内容简介

本书是“大学数学的内容、方法与技巧丛书”之一，是大学生学习概率论与数理统计的优秀辅导书和报考研究生的必备参考书，更是有志于掌握概率论与数理统计方法的读者的一本极好的指导书.

本书从教育部关于《概率论与数理统计课程的教学要求》与《硕士研究生入学考试数学考试大纲》出发，并略有提高地按章节对各个问题的内容、方法与技巧进行了归纳提高、释疑解难、分析演绎，以帮助读者理解和掌握概率论与数理统计方法.

本书内容包括随机事件与概率、随机变量及其概率分布、多微随机变量及其分布、随机变量的数字特征、大数定律与中心极限定理、数理统计的基本概念、参数估计、假设检验、方差分析与回归分析等，还附有实行全国硕士研究生入学统一考试以来的概率论与数理统计试题的解答，提供给考研读者作为参考.

希望本书能成为读者的良师益友，欢迎读者选用本系列丛书.

第二版前言

“概率论与数理统计”是高等学校的一门重要的数学基础课，也是考研数学的重要组成部分，概率统计方法更是科学技术、经济管理、工农业生产和社会人文各个领域中卓有成效的处理问题、解决问题的方法. 广大读者需要概率论与数理统计，喜爱概率论与数理统计，但也感到概率论与数理统计的概念难懂、方法难以掌握、思维难以展开、问题难以入手和习题难做. 本书从读者的角度出发，帮助大家解决学习中的种种困难.

本书按照教育部关于《概率论与数理统计课程的教学要求》和《硕士研究生入学考试数学考试大纲》编写，并在此基础上略有提高. 因此，特别适合在校大学生和有志于报考硕士研究生的人士使用.

为了使读者能够循序渐进、扎扎实实地从理论上、方法上和实践上掌握概率论与数理统计的概念与方法，我们采取以章节为序的方法，逐个问题地进行讨论、分析、讲解、举例、演绎、归纳. 每一节先对概念、内容进行梳理、归纳、提炼，然后对内容、方法中问题进行讨论、释疑解难，再对方法、技巧进行典型例题分析，边演绎、边讨论、边总结，最终达到消化、理解和掌握的目的. 为此，作者用解析方法认真地对读者学习中可能产生的对概念的误解、对方法的错失进行了分析探讨、论证求索，选用了较全面、较典型的例题帮助读者理解概率统计的思想方法、步骤和最终的结论. 相信本书能给读者以启迪和帮助，使读者能更好掌握概率统计方法.

本书第一版得到读者的厚爱，为答谢读者，这次在保持原有风格的基础上，在内容调整、例题演算、语言表达等方面进行了全面的修订，以使本书更加贴近读者.

为了帮助读者准备硕士研究生入学考试的数学考试，本书对1987年以来全国工学、经济学硕士研究生入学数学考试试卷中的概率统计试题作了全面和详尽的解答，并着意加强了最近的硕士研究生入学试题分析的内容，与考研数学要求一起附在每章的后面. 读者可以从中了解考研的要求、考点与动向.

本书在编写、出版过程中，得到华中科技大学出版社的热心支持与帮助，在此表示衷心的感谢.

对于本书中可能出现的错漏和失误之处，热忱欢迎同行和读者给予批评、指正.

孙清华　孙　昊

2006年3月于武汉

目　录

第一章　随机事件与概率

第一节　样本空间与随机事件

主 要 内 容

1. 随机现象

在一次试验中可能出现不同结果，而在大量重复试验中各个结果呈现统计规律性的现象称为随机现象.

如，在正常条件下，水加热到100℃会沸腾，是确定性现象；足球运动员临门施射不一定能射中，是不确定性现象.

2. 随机试验

若把科学实验或观察都称为试验，则满足下列条件的试验称为随机试验：

(1) 在相同条件下可以重复进行；

(2) 每次试验的可能结果不止一个，且在试验开始前能明确所有可能的结果；

(3) 每次试验前不能确定哪个结果会出现.

随机试验一般用大写字母E，F，…来表示，我们通过随机试验来研究随机现象.

3. 样本空间

随机试验的每一个可能出现的不可分解的结果称为样本点，全体样本点的集合称为样本空间，用Ω(或S)来表示.

4. 随机事件

样本空间Ω的子集合称为试验E的随机事件，简称事件，以大写字母$A,B,\cdots$来表示. 随机事件可以分为：

(1) 基本事件　只含一个样本点的子集合.

(2) 复合事件　含若干个样本点的子集合.

(3) 不可能事件　不含样本点的子集合(空集)，所以它在每次试验中都不会发生，记为$\varnothing$.

(4) 必然事件　样本空间本身，所以它在每次试验中必然发生，记为Ω.

事实上，(3)与(4)具有确定性，不是随机事件，但仍可把它们当作随机事件来处理.

5. 事件的关系

设Ω为试验E的样本空间，A,B,C为Ω的子集，则以下关系存在：

(1) 包含　若A的每个样本点都属于B，则A发生导致B发生，称事件B包含事件A，或事件A被事件B包含，记为$A\subset B$.

(2) 等价　若$A\supset B$与$B\supset A$同时成立，则称A与B等价，记为$A=B$. 在一次试验中，等价的两个事件或同时发生或同时不发生.

(3) 互斥(互不相容)　若事件A与事件B不能同时发生，称事件A与B互不相容(互斥)，记为$A\cap B=\varnothing$(或$AB=\varnothing$).

6. 事件的运算

由于事件是集合，因此事件的运算与集合的运算是一致的. 常用的运算如下：

(1) 并(和)　至少属于A或B中一个的所有样本点的集合称为事件A与B的并(或和)，记为$A+B$或$A\cup B$. 即在一次试验中，$A+B$发生表示A与B至少有一个发生.

$A_1+A_2+\cdots+A_n$或$\bigcup\limits_{k=1}^{n}A_k$称为$n$个事件$A_1,A_2,\cdots,A_n$的和. $A_1+A_2+\cdots+A_n+\cdots$或$\bigcup\limits_{k=1}^{\infty}A_k$称为可列个事件$A_1,A_2,\cdots,A_n,\cdots$的和.

(2) 交(积)　同时属于 A 和 B 的所有样本点的集合称为事件 A 与 B 的交(或积),记为 $A\cap B$ 或 AB. 在一次试验中,AB 发生表示 A 与 B 都发生.

$A_1A_2\cdots A_n$ 或 $\bigcap_{k=1}^{n}A_k$ 称为 n 个事件 $A_1,A_2,\cdots,A_n$ 的积. $A_1A_2\cdots A_n\cdots$ 或 $\bigcap_{k=1}^{\infty}A_k$ 称为可列个事件 $A_1,A_2,\cdots,A_n,\cdots$ 的积.

(3) 差　事件 A 发生而事件 B 不发生,称为 A 与 B 的差,记为 $A-B$,有关系式 $A-B=A\overline{B}$.

需要注意的是,不要求 $A\supset B$ 才有 $A-B$,如,图 1.1 阴影部分即为 $A-B$.

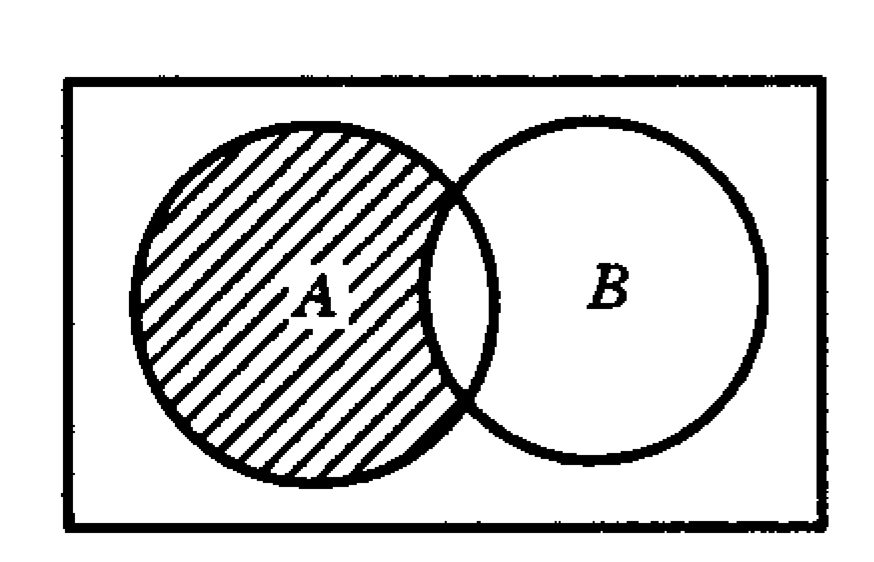


图 1.1

(4) 逆(对立)　样本空间 Ω 中所有不包含在 A 中的样本点的集合称为 A 的逆,记为 $\overline{A}$,也称为 A 的对立事件. 在一次试验中 $\overline{A}$ 发生表示 A 不发生. 有关系式

$$A+\overline{A}=\Omega,\quad A\cap\overline{A}=\varnothing.$$

7. 事件的运算规律

(1) 交换律　$A\cup B=B\cup A, A\cap B=B\cap A$.

(2) 结合律　$A\cup(B\cup C)=(A\cup B)\cup C$,

$A\cap(B\cap C)=(A\cap B)\cap C$.

(3) 分配律　$A\cap(B\cup C)=(A\cap B)\cup(A\cap C)$,

$A\cup(B\cap C)=(A\cup B)\cap(A\cup C)$.

(4) 对偶原理　$\overline{A\cup B}=\overline{A}\cap\overline{B}, \overline{A\cap B}=\overline{A}\cup\overline{B}$;

$A+A=A,\ A+\Omega=\Omega;\quad A\Omega=A,\ A\varnothing=\varnothing$.

疑 难 解 析

1. 怎样确定随机试验的样本空间?

答　对一个随机试验而言,样本空间并不一定唯一. 在同一试验中,当试验的目的不同时,样本空间往往是不同的. 如,把篮球运

动员投篮作为随机试验时，若试验目的是考察命中率，则试验的样本空间为$\Omega_1=\{$中，不中$\}$；若试验目的是考察得分情况，则试验的样本空间为$\Omega_2=\{1$ 分，2 分，3 分，0 分$\}$. Ω_1 与 Ω_2 显然不同. 所以，我们应从试验目的出发来确定样本空间.

2. 怎样理解样本空间与必然事件的关系？

答 必然事件与样本空间的关系应当这样来认识：必然事件是指随机试验中一定会出现的事件. 当在一次试验中只有一个样本点出现时，如果把样本空间视作一个整体，就可以说样本空间 Ω 在每次试验中都出现了. 因而样本空间是随机试验的必然事件.

3. 如何认识互逆事件与互斥事件之间的联系与区别？

答 A 与 B 互逆，则$B=\overline{A}$. 在一次试验中，A 与 B 必有一个发生，且至多只有一个发生.

如果事件 A 与 B 不能同时发生，则 A，B 互斥. 但 A，B 也可以同时不发生. 因此，互逆必定互斥，互斥不一定互逆.

区别互逆与互斥的关键是：互逆只在样本空间只有两个（或两类）事件时存在，互斥还可在样本空间有多个（或多类）事件时存在. 互斥事件的特征是：在一次试验中，两个互斥事件可以同时不发生. 如，在一次考试中，及格与不及格总有一个发生，它们互逆又互斥；但考试成绩为70 分或80 分是互斥的，却不互逆，因为它们可以同时不发生.

4. 随机事件的运算与数的运算是否相同？

答 不相同. 不能把随机事件的“积”与“和”理解成数的“积”与“和”. 虽然它们性质类似，都满足交换律、结合律和分配律，但不能认为它们就是相同的运算. 事实上，它们是完全不同的运算，反映了不同的两类概念. 如：

对于数 a，有 $a+a=2a$，$aa=a^2$；而对于事件 A，有

$$A+A=A,\quad AA=A.$$

对于数 a，b，c，有 $a+bc\neq(a+b)(a+c)$；而对于事件 A，B，C，有 $A+BC=(A+B)(A+C)$.

方法、技巧与典型例题分析

本节常见的习题类型是：用简单事件表示复合事件、证明关于事件的等式或不等式、用事件表示应用问题的结果等.

常用的方法是：(1) 利用运算性质与规律将复合事件用等价的简单事件表示；(2) 利用集合的文氏图分析事件间关系，找出等价事件.

例 1 设 A,B 为任意两个事件，则

$$(\overline{A}+B)(A+B)(\overline{A}+\overline{B})(A+\overline{B})=\underline{\qquad\qquad}.$$

解 因为 A 与 $\overline{A}$ 互逆，B 与 $\overline{B}$ 互逆，所以

$$\begin{aligned}(\overline{A}+B)(A+B)&=\overline{A}A+\overline{A}B+BA+BB\\&=\varnothing+B+B=B,\\(\overline{A}+\overline{B})(A+\overline{B})&=\overline{A}A+\overline{A}\overline{B}+\overline{B}A+\overline{B}\overline{B}\\&=\varnothing+\overline{B}+\overline{B}=\overline{B},\end{aligned}$$

于是 $$(\overline{A}+B)(A+B)(\overline{A}+\overline{B})(A+\overline{B})=B\overline{B}=\varnothing.$$

例 2 设 A,B 为任意两个事件，则 $(A+B)(\overline{A}+\overline{B})$ 表示 ( ).

(A) 必然事件；　　(B) A 与 B 恰有一个发生；

(C) 不可能事件；　　(D) A 与 B 不同时发生.

解 选(B). 因为 $\overline{A}+\overline{B}=\Omega-AB$，所以

$$\begin{aligned}(A+B)(\overline{A}+\overline{B})&=(A+B)(\Omega-AB)=A\Omega-AB+B\Omega-AB\\&=A+B-AB,\end{aligned}$$

表明 A 与 B 恰有一个发生.

例 3 对任意两事件 A,B，证明：$A-B=A\overline{B}$.

证 设 $x\in(A-B)$，则 $x\in A$，$x\overline{\in}B$，即 $x\in A\overline{B}$，从而 $A-B\subset A\overline{B}$. 反之，若 $x\in A\overline{B}$，则 $x\in A$，$x\overline{\in}B$，即 $x\in(A-B)$，从而 $A\overline{B}\subset A-B$. 于是，$A-B=A\overline{B}$ 得证.

例 4 指出下列各式成立的条件：

(1) $A-(B-C)=(A-B)+C$;

(2) $(A+B)-C=A+(B-C)$;

(3) $ABC=AB(C+B)$;

(4) $\overline{(A+B)}C=C-C(A+B)$.

解 (1) 由例3知

$$A-(B-C)=A\,\overline{(B\,\overline{C})}=A(\overline{B}+\overline{\overline{C}})=A\overline{B}+AC,$$

而$A-B=A\overline{B}$,故要$A-(B-C)=(A-B)+C$,必须有$(A-B)+C=A\overline{B}+AC$,即$AC=C$.从而知C是A的子集.

由此可知,在代数运算中的去括号与消去律在事件运算中是不成立的,一定要牢记两种运算的区别.

(2) 将式子变形为$(A+B)\overline{C}=A+B\overline{C}$,得$A\overline{C}+B\overline{C}=A+B\overline{C}$,知$AC=\varnothing$时等式成立.

(3) 将式子变形为$AB(C+B)=ABC+AB=ABC$,得$AB=ABC$,知$C\supset AB$时等式成立.

(4) $\overline{(A+B)}C=[\Omega-(A+B)]C=C-(A+B)C=C-C(A+B)$,知等式恒成立.

例5 利用事件间关系,化简下列式子:

(1) $(A+B)(A+C)$; (2) $\overline{(\overline{A}\overline{B}+C)\overline{A}\overline{C}}$;

(3) $(A+B)(A+\overline{B})$.

解 (1) $(A+B)(A+C)=AA+AB+AC+BC=A+AB+AC+BC$,而$(AB+AC)\subset A$,故

$$(A+B)(A+C)=A+BC.$$

(2) 本题中含有多个逆事件,因此要多次使用对偶律,以除去“逆”记号.

$$\begin{aligned}\overline{(\overline{A}\overline{B}+C)\overline{A}\overline{C}}&=\overline{(\overline{A}\overline{B}+C)}+\overline{\overline{A}\overline{C}}=AB\overline{C}+AC\\&=(A+B)\overline{C}+AC=A\overline{C}+B\overline{C}+AC\\&=A(C+\overline{C})+B\overline{C}=A+B\overline{C}.\end{aligned}$$

(3) $$\begin{aligned}(A+B)(A+\overline{B})&=(A+B)A+(A+B)\overline{B}\\&=A+AB+A\overline{B}+B\overline{B}\end{aligned}$$

$$=A+A(B+\overline{B})+\varnothing=A.$$

例6 证明下列等式：

(1) $A\cup B=A\cup B\overline{A}$；　　(2) $B-A=\overline{AB}-\overline{\overline{A}B}$；

(3) $(A-B)\cup(B-A)=\overline{AB\cup\overline{A}\,\overline{B}}$.

证 利用运算性质和运算律，有

(1) $A\cup B=(A\cup B)\cap\Omega=(A\cup B)\cap(A\cup\overline{A})$

$=A\cup AB\cup B\overline{A}=A\cup B\overline{A}$.

(2) $\overline{AB}-\overline{\overline{A}B}=\overline{AB}\cap\overline{A}B=(\overline{A}\cup\overline{B})\cap(\overline{A}B)=\overline{A}B$.

(3) $\overline{AB\cup\overline{A}\,\overline{B}}=\overline{AB}\cap\overline{\overline{A}\,\overline{B}}=(\overline{A}\cup\overline{B})\cap(A\cup B)$

$=A\overline{A}\cup\overline{B}A\cup\overline{A}B\cup\overline{B}B=A\overline{B}\cup B\overline{A}$

$=(A-B)\cup(B-A)$.

例7 用文氏图说明下列各式：

(1) $(A\cup B)C=AC\cup BC$；

(2) $AB\cup C=(A\cup C)(B\cup C)$；

(3) $A+B+C$.

解 (1) 如图1.2(a)表示，$(A\cup B)C$ 表示 A 与 B 之和与 C 的交集，是画有交叉线的阴影部分；AC 是 A 与 C 之交，是画有右斜线的阴影部分，BC 是 B 与 C 之交，是画有左斜线的阴影部分. AC 与 BC 之积是画有交叉线的阴影部分，故知两者相等.

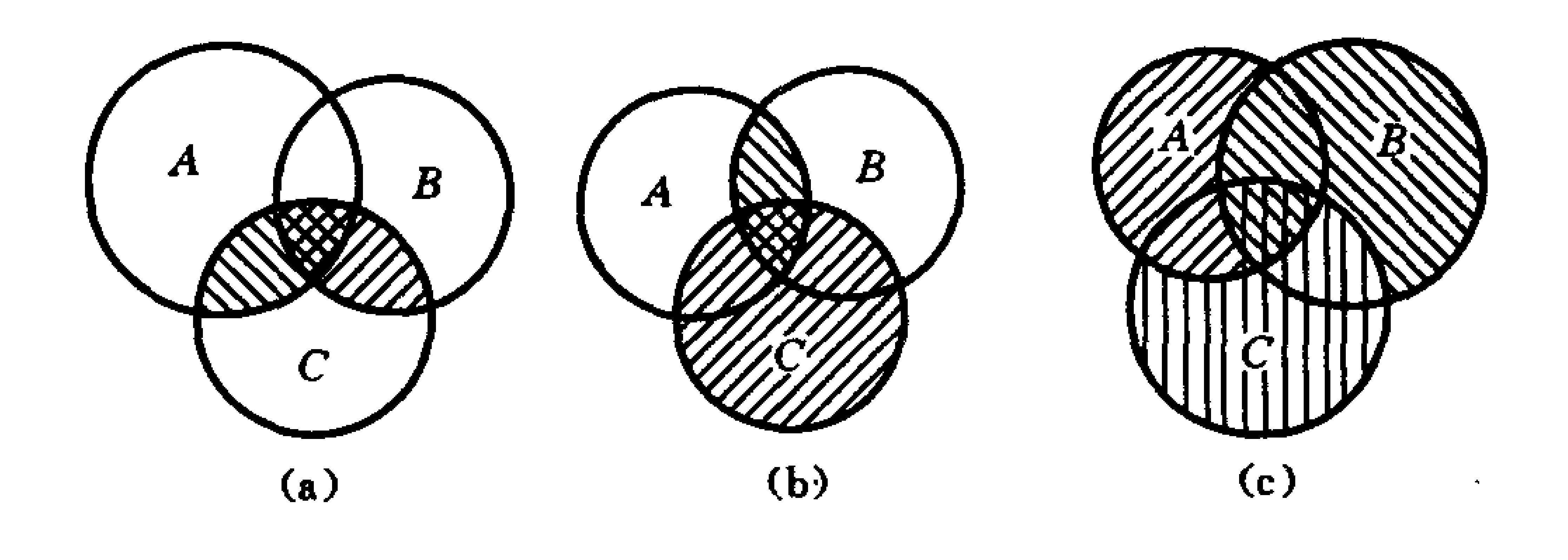


图1.2

(2) $AB\cup C$ 表示 A 与 B 之和与 C 的交集，是图1.2(b)中画有交叉线的部分；$(A\cup C)(B\cup C)$表示 A 与 C 之和同 B 与 C 之和的

交，两者是相等的.

(3) $A+B+C$ 是 A,B,C 的和集(见图 1.2(c))，可以表示为

$$A+B+C=(A-AB)+(B-BC)+(C-AC)+ABC.$$

这里，因为 $(B-BC)$ 与 $(C-CA)$ 中重复减去了 ABC，所以最后要加上一个 ABC.

例8 如果以 x 表示一个沿数轴作随机运动的质点的位置，说明下列事件的关系：

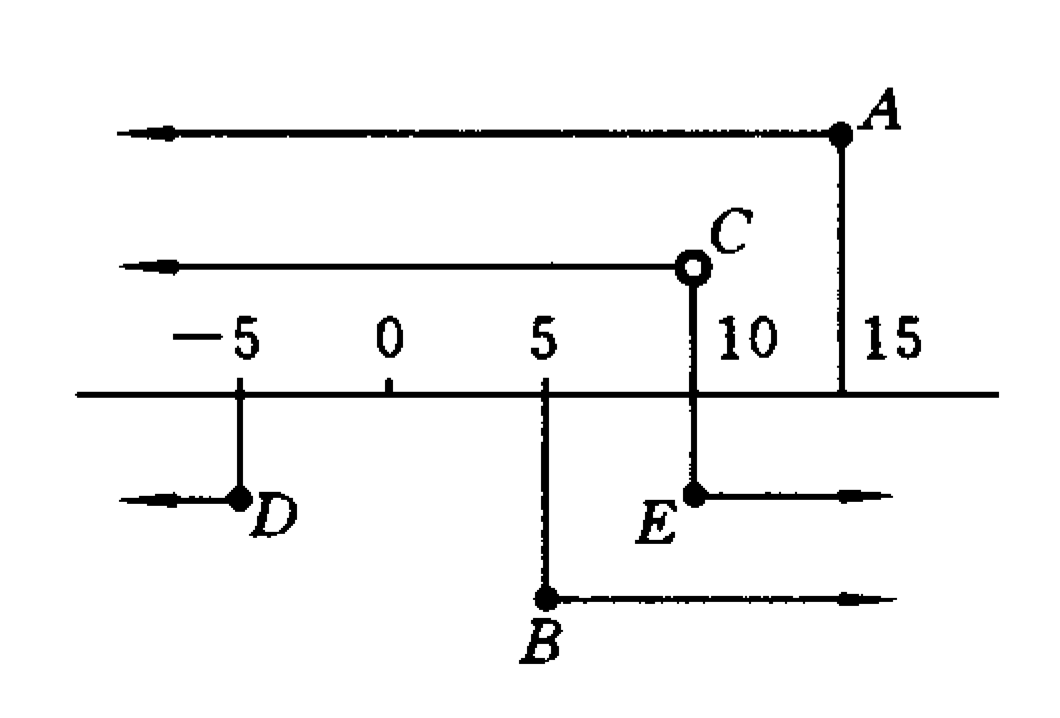


图 1.3

$$A=\{x\,|\,x\leqslant 15\},$$
$$B=\{x\,|\,x\geqslant 5\},$$
$$C=\{x\,|\,x<10\},$$
$$D=\{x\,|\,x\leqslant -5\},$$
$$E=\{x\,|\,x\geqslant 10\}.$$

解 由图 1.3 可知：$A\supset C\supset D$；$B\supset E$；D 与 B 互斥，D 与 E 互斥；B 与 C 相容，B 与 A 相容，E 与 A 相容；C 与 E 互逆.

例9 射击运动员射击的目标是三个半径(单位：m)分别为 0.1，0.2，0.3 的同心圆环域，标为 r_1,r_2,r_3，以 A_i $(i=1,2,3)$ 记击中半径为 r_i 的圆环域内事件，试以事件的集合表示下列情况：

(1) 击中 0.3 m 半径的圆环域外；

(2) 击中任一圆环域内；

(3) 击中 0.1 m 半径的圆环域内；

(4) 击中 0.1 m 半径的圆环域外，0.2 m 半径的圆环域内.

解 (1) $\overline{A}_3$，即 A_3 的逆事件； (2) $A_1\cup A_2\cup A_3$；

(3) A_1； (4) A_2.

例10 一批产品中有合格品也有废品，从中有放回地抽取(将产品取出一件观察后放回)三件产品，以 A_i $(i=1,2,3)$ 表示第 i 次抽到废品，试以事件的集合表示下列情况：

(1) 第一次和第二次抽取至少抽到一件废品；

(2) 只有第一次抽到废品； (3) 三次都抽到废品；

(4) 至少有一次抽到废品；　　(5) 只有两次抽到废品.

解　(1) $A_1 \cup A_2$；　(2) $A_1 \cap \overline{A}_2 \cap \overline{A}_3$；　(3) $A_1A_2A_3$；

(4) $A_1 \cup A_2 \cup A_3$ 或 $\overline{A_1A_2A_3}$(三次都抽到合格品的逆事件)；

(5) $\overline{A}_1A_2A_3 \cup A_1\overline{A}_2A_3 \cup A_1A_2\overline{A}_3$.

例 11　设有随机事件 A,B,C，满足 $C \supset AB$，$\overline{C} \supset \overline{A}\overline{B}$，证明：

$$AC = C\overline{B} \cup AB.$$

证　因为 $\overline{C} \supset \overline{A}\overline{B}$，所以 $C \subset A \cup B$. 于是，$C\overline{B} \subset (A \cup B)\overline{B} = A\overline{B}$，$CA\overline{B} = C\overline{B} \cap A\overline{B} = C\overline{B}$，$ACB = C \cap AB = AB$. 故

$$\begin{aligned} AC &= AC(B \cup \overline{B}) = AC\overline{B} \cup ACB \\ &= C\overline{B} \cup AB. \end{aligned}$$

由文氏图(见图 1.4)可以清楚地看出，AC 为全部画有斜线的部分，$C\overline{B}$ 为画有左斜线的部分，AB 为画有右斜线的部分. 所以结论成立.

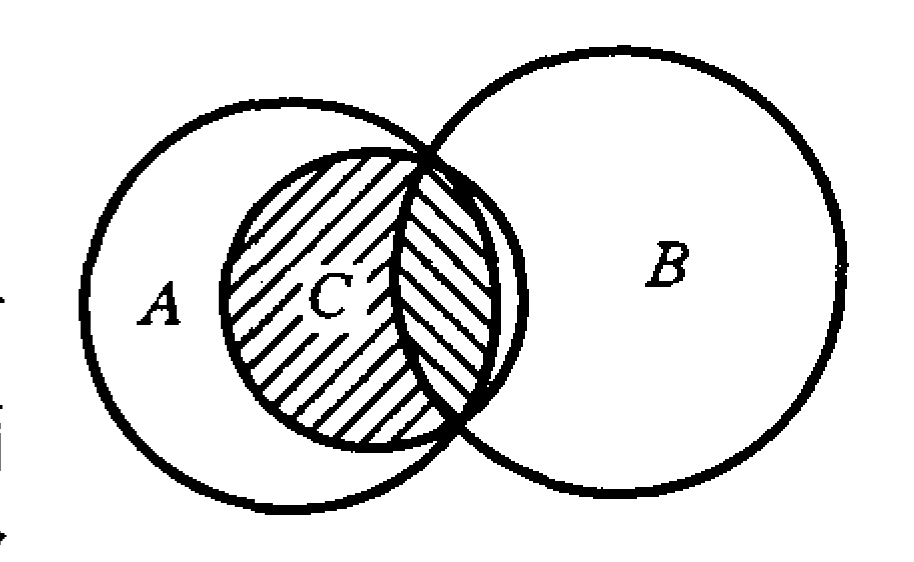


图 1.4

第二节　随机事件的概率

主要内容

1. 频率

在相同的条件下进行了 n 次试验，若事件 A 发生了 n_A 次，则 n_A 称为 n 次试验中 A 发生的频数. $f_n(A) = n_A/n$ 称为事件 A 发生的频率. 频率具有以下性质：

(1) 对任何事件 A，有 $0 \leqslant f_n(A) \leqslant 1$；

(2) 对必然事件 Ω，有 $f_n(\Omega) = 1$；

(3) 对 k 个两两互不相容事件 $A_1, A_2, \cdots, A_k$，有

$$f_n(A_1)+f_n(A_2)+\cdots+f_n(A_k)=f_n(\bigcup_{i=1}^{k}A_i)=\sum_{i=1}^{k}f_n(A_i).$$

2. 概率的公理化定义

设Ω是随机试验E的样本空间，对E的每一事件A赋予一个实数值，称为事件A的概率，记为$P(A)$. 函数$P(\cdot)$具有以下性质：

(1) 非负性　对任一事件A，$P(A)\geqslant 0$；

(2) 规范性　$P(\Omega)=1$；

(3) 可加性　对可列个两两互不相容事件$A_1,A_2,\cdots,A_n,\cdots$，有

$$P(A_1\cup A_2\cup\cdots\cup A_n\cup\cdots)=P(\bigcup_{i=1}^{\infty}A_i)=\sum_{i=1}^{\infty}P(A_i).$$

3. 概率的统计定义

当$n\to\infty$时，n次独立重复试验的频率$f_n(A)=n_A/n$趋于稳定值$P(A)$，则当n很大时，$P(A)=p\approx n_A/n$.

4. 概率的几何意义

若随机试验E的可能结果有无限(不可列)个，每个基本事件发生的可能性相等，则当样本空间Ω与所求事件A都可以用几何量(长度L、面积S或体积V)来测度时，A发生的概率称为几何概率. 例如

$$P(A)=\frac{S(A)}{S(\Omega)}.$$

其中$S(A)$，$S(\Omega)$分别称为A，Ω的几何测度.

5. 概率的性质

(1) $P(\varnothing)=0$；

(2) 对任一事件A，$P(A)=1-P(\overline{A})$；

(3) 若$A_1,A_2,\cdots,A_n$两两互不相容，则

$$P(A_1\cup A_2\cup\cdots\cup A_n)=P(\bigcup_{i=1}^{n}A_i)=\sum_{i=1}^{n}P(A_i);$$

(4) 对两个事件A,B，若$A\subset B$，则

$$P(B-A)=P(B)-P(A),\quad P(B)\geqslant P(A).$$

6. 概率的加法公式

(1) 对于任意两个事件 A,B,有

$$P(A+B)=P(A)+P(B)-P(AB).$$

(2) 对于任意 n 个事件 $A_1,A_2,\cdots,A_n$,有

$$P(A_1+A_2+\cdots+A_n)$$

$$=\sum_{i=1}^{n}P(A_i)-\sum_{1\leqslant i<j\leqslant n}P(A_iA_j)+\sum_{1\leqslant i<j<k\leqslant n}P(A_iA_jA_k)$$

$$-\cdots+(-1)^{n-1}P(A_1A_2\cdots A_n).$$

7. 古典型概率

若随机试验的样本空间的元素为有限个(有限性),每个样本点发生的可能性相等(等可能性),则试验 E 的事件 A 发生的概率 $P(A)$ 称为古典型概率. 记 m 为事件 A 中所含基本事件数,n 为样本空间 Ω 中基本事件的总数,则计算公式为

$$P(A)=\frac{m}{n}.$$

利用古典型概率讨论事件概率的数学模型称为古典概型.

疑 难 解 析

1. 怎样理解统计概率?

答 随机事件 A 的频率 $f_n(A)$ 反映在 n 重独立重复试验中 A 发生的频繁程度. 当 n 不同或在不同组试验时,$f_n(A)$ 的值一般是不同的. 但当 n 充分大时,$f_n(A)$ 会在概率 $P(A)=p$ 的值左右徘徊. n 越大,徘徊的区间越小. 所以,可以在这小区间中取一值作为 p 的近似值,于是有

$$P(A)=p\approx n_A/n=f_n(A).$$

统计概率是一个近似值.

2. 能否将概率看作频率的极限?

答 不能. 这是因为:第一,统计概率仅对 n 次独立重复试验

而言（伯努利大数定律），并非所有情形都适用；第二，当 $n\to\infty$ 时，$f_n(A)$ 在 $P(A)$ 左右徘徊，与微积分中 $f(x)\to A$ 不同．用 ε-N 概念来解释就是：存在随机现象的偶然性，对于某个给定的 $\varepsilon>0$，可能找不到相应的 $N(\varepsilon)$，使当 $n>N$ 时，有 $|f_n(A)-P(A)|<\varepsilon$ 成立．

3．怎样确定随机事件的等可能性？

答　等可能性是古典型概率问题的两大假设之一，有了这两大假设，我们只需讨论样本空间 Ω 和事件 A 所包含的基本事件数，即可计算概率 $P(A)$．但所讨论问题是否符合等可能假设，一般不可能实际验证，而是根据人们长期形成的"对称性经验"作出的．如，将一枚均匀的硬币掷一次，正面向上和反面向上的机会是相等的；大小、形状和重量相同而颜色不同的小球装在同一小盒子中，每个球被摸到的可能性也是相等的．但是，一个产妇到医院生产，产下男婴和女婴的可能性一般不相等；投一次篮，投中和投不中的可能性一般也不相等．因此，等可能性的确定是人们根据对事物的长期认识而作出的，不是人为的．

4．怎样判断讨论的问题是排列问题还是组合问题？

答　在计算样本空间 Ω 和事件 A 所包含基本事件数时，事件数的多少与问题是排列还是组合有关．当事件的组成与顺序有关时，是排列问题；与顺序无关时，是组合问题．

如，某班级有40名学生，要选3人分别担任班长、学习委员、文体委员，有两种方案．一种是选出3人，由他们自行分工，这种方案与顺序无关，是组合问题，共有 C_{40}^3 种选法．另一种方案是分别选出班长（40种选法）、学习委员（39种选法）、文体委员（38种选法），与顺序有关，是排列问题，共有 P_{40}^3 种选法．第二种方案也可以表示为 $P_{40}^3=C_{40}^3\times 3!$，即选出三人后，再自由排列．

5．怎样确定几何型概率中几何量的测度？

答　在几何型概率中，所考虑的问题中若只有一个因素在变，则取一维几何量——长度作几何测度（见本节例26）；若有两个因素在变，则取二维几何量——面积作几何测度（见本节例27～30）；

若有三个因素变化，则取三维几何量——体积作几何测度(见本节例31).

方法、技巧与典型例题分析

一、基本的概率问题

基本的概率问题一般可以利用事件之间的关系与运算，运用概率的运算性质求解或证明. 读者对基本概念应有较深刻的理解，熟练掌握事件与概率的运算.

例1 某医院一天中接诊外科病人50人，内科病人50人，五官科病人50人. 设每位病人在一科室至多就诊一次，在病人总数中，在三个科室各就诊一次的占10%，只看外科的占20%，只看内科的占25%，只看五官科的占10%. 问：

(1) 一天共接诊多少个病人?

(2) 只看外科和内科的病人占病人总数的比例为多少?

解 以 A,B,C 分别表示外科、内科、五官科接诊病人的集合，由文氏图(见图1.5)知

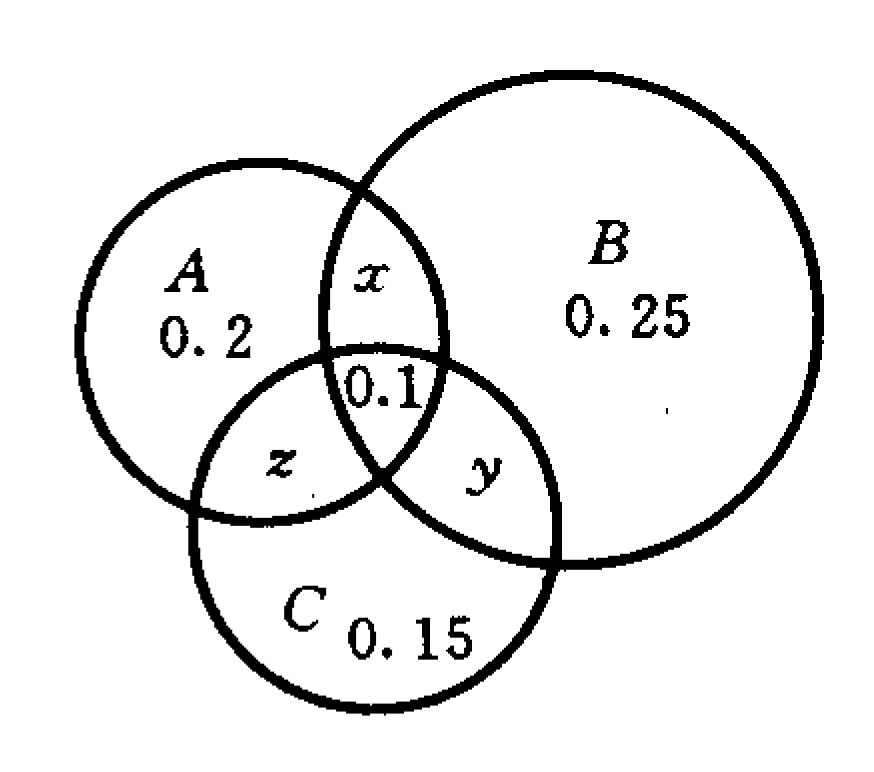


图1.5

$$A+B+C=ABC+A\overline{B}\overline{C}+\overline{A}B\overline{C}+\overline{A}\overline{B}C+\overline{A}BC+A\overline{B}C+AB\overline{C},$$

又知 $P(ABC)=0.1$， $P(A\overline{B}\overline{C})=0.2$，

$P(\overline{A}B\overline{C})=0.25$， $P(\overline{A}\overline{B}C)=0.15$.

设 $P(AB\overline{C})=x$， $P(\overline{A}BC)=y$，

$P(A\overline{B}C)=z$，又设病人总人数为 S，可建立方程组

$$\begin{cases} S(0.2+x+y+0.1)=50, \\ S(0.25+y+z+0.1)=50, \\ S(0.15+z+x+0.1)=50, \\ x+y+z=1-0.2-0.25-0.15-0.1. \end{cases}$$

解方程得

$$x=0.05,\quad y=0.15,\quad z=0.1,\quad S=100.$$

所以病人总数为 100 名，只看外科和内科的病人占总病人数的 5%.

例2 已知$P(A)=P(B)=P(C)=1/4, P(AB)=0, P(AC)=P(BC)=1/8$，则事件$A,B,C$全不发生的概率为________.

解 由$ABC\subset AB$知，$P(ABC)=0$. 又由逆事件公式和加法公式，得

$$\begin{aligned}P(\overline{A}\overline{B}\overline{C})&=1-P(A+B+C)\\&=1-P(A)-P(B)-P(C)+P(AB)+P(BC)\\&\quad+P(AC)-P(ABC)\\&=1-3\times1/4+2\times1/8=1/2.\end{aligned}$$

例3 设当事件A,B都发生时，事件C必发生，则(　　).

(A) $P(C)\leqslant P(A)+P(B)-1$；　　(B) $P(C)=P(AB)$；

(C) $P(C)\geqslant P(A)+P(B)-1$；　　(D) $P(C)=P(A+B)$.

解 由题意知$C\supset AB$，故

$$\begin{aligned}P(C)\geqslant P(AB)&=P(A)+P(B)-P(A+B)\\&\geqslant P(A)+P(B)-1,\end{aligned}$$

所以选(B).

例4 设$P(A)=a, P(B)=2a, P(C)=3a, P(AB)=P(BC)=b$，证明：$a\leqslant1/4$.

证 由$P(AB)\subset P(A)$，得$b\leqslant a$. 又由加法公式

$$\begin{aligned}1\geqslant P(B+C)&=P(B)+P(C)-P(BC)\\&=2a+3a-b\geqslant4a,\end{aligned}$$

所以
$$a\leqslant1/4.$$

例5 对任意事件$A_1,A_2,\cdots,A_n$，证明：

(1) $P(A_1A_2)\geqslant P(A_1)+P(A_2)-1$；

(2) $P(A_1A_2\cdots A_n)\geqslant P(A_1)+P(A_2)+\cdots+P(A_n)-(n-1)$.

证 (1) 由加法公式

$$P(A_1+A_2)=P(A_1)+P(A_2)-P(A_1A_2),$$

所以

$$P(A_1A_2)=P(A_1)+P(A_2)-P(A_1+A_2)$$
$$\geqslant P(A_1)+P(A_2)-1.$$

(2) 用数学归纳法. $n=2$ 即有题(1)的结论. 设对 $n-1$,不等式成立,则

$$\begin{aligned}&P(A_1A_2\cdots A_n)\\&=P[(A_1A_2\cdots A_{n-1})A_n]\\&\geqslant P(A_1A_2\cdots A_{n-1})+P(A_n)-1\\&\geqslant P(A_1)+P(A_2)+\cdots+P(A_{n-1})-(n-2)+P(A_n)-1\\&=P(A_1)+P(A_2)+\cdots+P(A_n)-(n-1).\end{aligned}$$

例6 设有甲、乙两人玩投篮游戏,规定:每轮由甲先投一次,接着乙可投两次,先投中者胜. 已知甲每次投篮命中率为 p,乙命中率为0.5,问 p 取何值时,甲、乙两人胜负概率相等.

解 以 A_i,B_i $(i=1,2,\cdots)$分别记甲、乙在第 i 次投篮命中事件,i 又为甲、乙两人投篮的总次数,以 A,B 分别记甲、乙取胜事件,则

$$A=A_1+\overline{A}_1\overline{B}_2\overline{B}_3A_4+\overline{A}_1\overline{B}_2\overline{B}_3\overline{A}_4\overline{B}_5\overline{B}_6A_7+\cdots,$$

而

$$P(A_1)=p,$$
$$P(\overline{A}_1\overline{B}_2\overline{B}_3A_4)=(1-p)\times 0.5^2p=0.25p(1-p),$$
$$P(\overline{A}_1\overline{B}_2\overline{B}_3\overline{A}_4\overline{B}_5\overline{B}_6A_7)=0.25^2(1-p)^2p,$$

所以

$$P(A)=p+0.25(1-p)p+0.25^2(1-p)^2p+\cdots$$
$$=p/[1-0.25(1-p)].$$

由题设,要 $P(A)=P(B)=0.5=p/[1-0.25(1-p)]$,解得 $p=3/4$. 即当甲的命中率为3/4时,甲、乙胜负的概率相等.

二、古典型概率问题

古典型概率问题有三大典型问题:摸球问题、质点入盒问题和随机取数问题. 另外还有一些其它类型的问题,如超几何分布概率问题等,也属于古典型概率问题.

求解古典型概率问题的常用方法有:

1. 加法原理

设完成一件事有 k 类方法，每类分别有 $m_1, m_2, \cdots, m_k$ 种方法，而完成这件事只要选择任一类方法中的任何一种，则完成这件事的方法有 $m_1+m_2+\cdots+m_k$ 种.

2. 乘法原理

设完成一件事有 n 个步骤：第一步有 m_1 种方法，第二步有 m_2 种方法……第 n 步有 m_n 种方法，且这 n 类的所有种方法都各不相同，则完成这件事的方法有 $m_1m_2\cdots m_n$ 种.

注意，如果完成一件事的方法是并列的关系，则使用加法原理；如果完成一件事的方法有顺序关系，则使用乘法原理.

3. 排列方法

从全部元素中取出一部分，有次序地排成一列，称为一个排列. 排列可分为：

(1) 不同元素的选排列　从 n 个不同的元素中无放回地取出 m $(m<n)$个元素的排列，共有 P_n^m 种排列，

$$P_n^m=n(n-1)\cdots(n-m+1).$$

当 $m=n$ 时，称为全排列，共有 $n!$ 种排列.

(2) 不同元素的重复排列　从 n 个不同的元素中有放回地取出 m 个元素的排列，共有 n^m 种排列.

(3) 不全相似元素的排列　若在 n 个元素中有 m 类不同的元素，每类各有 $k_1, k_2, \cdots, k_m$ 个(每类元素彼此视为不可辨认)，则这 n 个元素的全排列共有 $n!/(k_1!k_2!\cdots k_m!)$种.

4. 环排列

从 n 个不同的元素中，选出 m 个元素排成一个圆周的排列，共有 $C_n^m(m-1)!$ 种排列.

5. 组合方法

从全部元素中取出一部分元素而不考虑其顺序的排列称为一个组合. 组合可分为：

(1) 一般的组合　从 n 个不同元素中取出 m 个的组合，共有

$$\frac{n(n-1)\cdots(n-m+1)}{m!}=\frac{n!}{m!\ (n-m)!}$$

种，也记为C_n^m或$\binom{n}{m}$. 它具有性质

$$C_n^m=C_n^{n-m},\quad C_n^m=C_{n-1}^m+C_{n-1}^{m-1},\quad C_n^0=1.$$

（2）不同类元素的组合　从不同的k类元素中取出m个元素，即从第一类的n_1个不同元素中取m_1个，从第二类的n_2个不同元素中取m_2个……从第k类的n_k个不同元素中取m_k个，且$n_i\geqslant m_i$ $(i=1,2,\cdots,k)$，$\sum_{i=1}^{k}m_i=m$，则共有组合$C_{n_1}^{m_1}C_{n_2}^{m_2}\cdots C_{n_k}^{m_k}=\prod_{i=1}^{k}C_{n_i}^{m_i}$种.

在处理具体问题时，既要考虑事件是否与顺序有关，从而确定用组合方法还是排列方法，又要从实际问题出发判断是哪一种排列或组合，才能求得正确的结果.

摸球问题是指从n个可分辨的球中按照不同的要求逐个地取出m个球，并计算事件概率的问题.

例7　一袋内有a个白球，b个黑球，求：

（1）不放回地任取$m+n$个球，恰有m个白球、n个黑球的概率；

（2）不放回抽取，每次一个，第k次才取到白球的概率；

（3）不放回抽取，每次一个，第k次恰取到白球的概率.

解　（1）任取$m+n$个球，与次序无关，且m个白球从a个白球中选，n个黑球从b个黑球中选，都是组合问题. 所以

$$p=\frac{C_a^mC_b^n}{C_{a+b}^{m+n}}\quad（属于超几何分布）.$$

（2）抽取与次序有关，第k次抽取的白球可为a个白球中任一个，所以

$$p=\frac{b}{a+b}\cdot\frac{b-1}{a+b-1}\cdot\cdots\cdot\frac{b-k+2}{a+b-k+2}\cdot\frac{C_a^1}{a+b-k+1}.$$

（3）第k次必取到白球，可为a个白球中任一个. 前$k-1$次取到的白球数应小于a个，因此第一次只能在$a+b-1$个球中取（留

一个白球在第k次取)，以后各次相同.所以

$$p=\frac{a+b-1}{a+b}\cdot\frac{a+b-2}{a+b-1}\cdot\cdots\cdot\frac{a+b-k+1}{a+b-k+2}\cdot\frac{C_a^1}{a+b-k+1}$$
$$=\frac{a}{a+b}.$$

例8 有n双不同的鞋混放在一起，有n个人每人随机地取走2只，求下列事件的概率：

(1) 每人取走的鞋恰为一双的概率；

(2) 每人取走的鞋不成一双的概率.

解 设$2n$只鞋被n个人取走.第一个人从$2n$只中任取2只，第二个人从$2n-2$只中任取2只……第n个人取走最后2只，但每个人取走2只的排列只是一种.所以，依乘法原理，基本事件的总数为

$$\frac{2n(2n-1)}{2}\cdot\frac{(2n-2)(2n-3)}{2}\cdot\cdots\cdot\frac{2\cdot 1}{2}=\frac{(2n)!}{2^n}.$$

(1) 每人取走的鞋恰为一双的事件数为$C_n^1C_{n-1}^1\cdots C_2^1C_1^1=n!$，于是

$$p=n!/[(2n)!/2^n]=(2n)!!/(2n)!.$$

(2) 每人取走的2只鞋都不成双的事件数为$(n!)^2$.因为第一个人可以从n只右脚鞋中取1只，又可以从n只左脚鞋中取1只(只要两只鞋不成一双)，其余依此类推.于是

$$p=(n!)^2/[(2n)!/2^n]=2^n(n!)^2/(2n)!\ =n!/(2n-1)!!.$$

例9 从5双不同的鞋子中任取4只，求这4只鞋子中至少有2只鞋子配成一双的概率.

解 本题的解法很多，我们仅举几种解法供读者参考.读者可尝试其它合理的解法.

(1) 基本事件总数为C_{10}^4.有利事件数包括：恰有2只配成一双，事件数是$C_5^1C_4^2C_2^1C_2^1$(C_5^1表示5双中任取1双，其余2只由4双中取2双，再在每1双中取1只)；4只配成两双，事件数是C_5^2.所求概率为

$$p=(C_5^1C_4^2C_2^1C_2^1+C_5^2)/C_{10}^4=130/210=13/21.$$

(2) 用逆事件求，4只鞋配不成双的事件数是$C_5^4C_2^1C_2^1C_2^1C_2^1$，故

$$p=1-(C_5^4C_2^1C_2^1C_2^1C_2^1)/C_{10}^4=1-80/210=13/21.$$

(3) 至少 2 只能配成一双也可以这样构成：先从 5 双中任取 1 只，其余 2 只从余下 8 只中取，但要减去可能成双的重复事件数，则

$$p=(C_5^1C_8^2-C_5^2)/C_{10}^4=13/21.$$

本题最容易出现的错误就是把有利事件数取为 $C_5^1C_8^2$，从而出现重复事件. 这是因为，若鞋子标有号码 1，2，…，5，则 C_5^1 可能取中第 i 号鞋，C_8^2 可能取中第 j 号的一双，此时成为两双的配对为 (i,j)，但也存在配对 (j,i). (i,j) 与 (j,i) 是一种，出现了重复事件，即多出了 $C_5^2=10$ 个事件.

例10 50 只铆钉随机地取来用在 10 个部件上，其中恰有 3 只铆钉强度太弱. 每个部件用 3 只铆钉，若将 3 只铆钉都装在同一部件上，则这个部件的强度就太弱. 问：发生一个部件强度太弱的概率是多大？

解 基本事件总数为 C_{50}^3. 强度太弱事件数为 $C_3^3C_{10}^1$（C_3^3 是强度太弱的铆钉组合，C_{10}^1是部件的取法数），所以

$$p=C_3^3C_{10}^1/C_{50}^3=1/1960.$$

例11 一袋内装9 个白球和3 个红球，从袋中任意地顺次取出 3 球（取出后不放回），求：

(1) 第三次取出的是白球的概率；

(2) 若第三次取出的是白球，则第一次取得的也是白球的概率.

解 (1) 同例 7，有 $p=9/12=3/4$.

(2) 第三次和第一次都取得白球的事件是两个互不相容事件：(白，白，白)，(白，红，白)，所以

$$p=\frac{9}{12}\times\frac{8}{11}\times\frac{7}{10}+\frac{9}{12}\times\frac{3}{11}\times\frac{8}{10}=\frac{72}{132}=\frac{6}{11}.$$

质点入盒问题是：有 n 个可分辨的盒子和 m 个质点，按照不同的要求将 m 个质点放在 n 个盒子中，计算事件的概率.

例12 将n 个质点随机地放入N $(N\geqslant n)$个盒子中，求下列事

件的概率：

(1) 某指定的n个盒子中各有一个质点；

(2) 任意n个盒子中各有一个质点；

(3) 指定的某盒中恰有m $(m<n)$个质点.

解 每个质点有N种放法，n个质点有N^n种放法，于是：

(1) 相当于n个质点在n个盒子中的全排列，共有$n!$种，所以

$$p=n!/N^n.$$

(2) 有利事件多了一个步骤，即n个盒子的选法有C_N^n种，所以

$$p=C_N^n n!/N^n.$$

(3) 从n个质点中取m个的取法有C_n^m种，其余的每个质点都有$N-1$种放法，所以

$$p=C_n^m(N-1)^{n-m}/N^n.$$

例13 某单位新录用了12名公务员，其中有3名博士. 将他们随机地平均分到三个研究室去，问：(1) 每一个研究室分到一名博士的概率是多少？(2) 3名博士分到同一研究室的概率是多少？

解 由不全相异元素的排列公式，得基本事件总数为

$$C_{12}^4C_8^4C_4^4=12!/(4!)^3.$$

(1) 将3名博士平均分配的分法有P_3^3种. 其余9名人员的分法有$9!/(3!)^3$种，有利事件数为$3!9!/(3!)^3$，所以

$$p=\frac{3!9!}{(3!)^3}\bigg/\frac{12!}{(4!)^3}=\frac{16}{55}.$$

(2) 将3名博士分到同一研究室的分法有C_3^1种，其余9人的分法有$9!/(1!4!4!)$种，有利事件数为$(3\times 9!)/(4!)^2$，所以

$$p=\frac{3\times 9!}{(4!)^2}\bigg/\frac{12!}{(4!)^3}=\frac{3}{55}.$$

例14 将3个球随机地放在4个杯子中去，求杯子中球的最大个数分别为1,2,3的概率.

解 基本事件总数为4^3个.

最大个数为1，含事件数为P_4^3. 所以，$p=P_4^3/4^3=3/8$.

最大个数为2，则有2个球的杯子的选法有C_4^1种，球的选法有C_3^2种，剩下1个球有3种放法，有利事件数为$3C_4^1C_3^2$. 所以

$$p=3C_4^1C_3^2/4^3=9/16.$$

最大个数为3，则杯子的选法为C_4^1，有利事件数为C_4^1. 所以

$$p=C_4^1/4^3=1/16.$$

例15 某班有12名学生是在1980年出生的，试求下列事件的概率：

(1) 至少有两人是同一天出生的；

(2) 至少有一人是5月1日出生的.

解 (1) 用求逆事件方法求解. 样本空间基本事件总数为365^{12}，没有两人在同一天出生的事件数为P_{365}^{12}，所以

$$p=1-P_{365}^{12}/365^{12}.$$

(2) 也用求逆事件方法求解. 没有人在5月1日出生的事件数为364^{12}，所以

$$p=1-364^{12}/365^{12}.$$

随机取数问题是：在n个不同的数中按不同的要求取出m个数，计算事件概率.

例16 随机地取一整数，求它的二次方的个位数字是4的概率.

解 设随机数$x=a+10b+\cdots$，其个位数是a，则$x^2=a^2+20ab+100b^2+\cdots$. 显然$x^2$的个位数仅与$x$的个位数$a$有关.

a有$0,1,\cdots,9$等10种取法，而使$a^2=4$的只有2种，$a=2$或$a=8$，故$p=2/10=1/5$.

例17 从$0,1,\cdots,9$等10个数字中，任取4个数字排成一排，求能成为四位偶数的概率.

解 排成四位偶数，显然与数字次序有关，基本事件总数为P_{10}^4. 要排成四位偶数，那么千位上不能取0，个位上只能取偶数，有$C_5^1P_9^3-P_8^2C_4^1$（前项是个位在5个偶数中选1个，另三位数从余下9

个数取3个，后项是千位为0时，个位有C_4^1种选法，其余两位有C_8^2种选法)．所以

$$p=(C_5^1P_9^3-C_4^1P_8^2)/P_{10}^4.$$

例18 从0,1,…,9等10个数字中任取一个，取出后仍放回，先后共取3次，求取出的三数总和等于15的概率.

解 基本事件总数为10^3. 每个数被取到的概率为1/10(等可能).

以a_i记第i次取到的数字，记A:$a_1+a_2+a_3=15$，则A所包含的基本事件相当于$(1+x+x^2+\cdots+x^9)^3$展开式x^{15}的系数. 因为由

$$\begin{aligned}&(1+x+x^2+\cdots+x^9)^3\\&=[(1-x^{10})/(1-x)]^3=(1-x^{10})^3(1-x)^{-3}\\&=(1-3x^{10}+\cdots)(1+\cdots+21x^5+\cdots+136x^{15}+\cdots)\end{aligned}$$

知，x^{15}的系数为$136-3\times21=73$，所以$p=73/10^3=0.073$.

例19 从1,2,…,20中任取一个数，取到数k的概率与k成正比，求取到的数是3的倍数的概率.

解 以p_k记取到数k的概率，则$p_k=ck$，且$\sum\limits_{k=1}^{20}ck=1$，可得$c=1/210$. 取到的数是3的倍数包括3,6,9,12,15,18，所以

$$p=(3+6+9+12+15+18)/210=3/10.$$

例20 将一枚硬币掷$2n$次，求出现正面次数多于反面次数的概率.

解 掷$2n$次硬币，可能出现：A——正面次数多于反面次数，B——反面次数多于正面次数，C——正面次数等于反面次数. A，B，C互不相容.

直接计算$P(A)$是困难的，可以用对称性原理来求解. 因为硬币是均匀的，有$P(A)=P(B)$，所以$P(A)=[1-P(C)]/2$. 而$2n$次掷硬币的排列有2^{2n}种，其中有利C的有C_{2n}^n种($2n$个中选n个)，则$P(C)=C_{2n}^n/2^{2n}$. 所以

$$P(A)=[1-P(C)]/2=(1-C_{2n}^n/2^{2n})/2.$$

例21 有外表相同的N ($N\geqslant 3$)个袋子，第k个袋子中装有k个红球和$N-k$个白球. 将袋的次序搞混后，任选一袋并任取一袋，求：

(1) 第一次取得红球、放回后第二次又取得红球的概率；

(2) 前两次取出的球不放回、第三次时取得红球的概率.

解 (1) 选中第k个袋子的概率是$1/N$，每次取得红球的概率是k/N，则由加法原理，有

$$p=\sum_{k=1}^{N}\frac{1}{N}\left(\frac{k}{N}\right)^2=\frac{1}{N^3}\sum_{k=1}^{N}k^2=\frac{(N+1)(2N+1)}{6N^2}.$$

(2) 选中第k袋的概率是$1/N$. 三次取球的方法有$N(N-1)(N-2)$种. 三次取得红球的取法有$k(k-1)(k-2)$种，取得(红，白，红)和(白，红，红)的取法各有$k(k-1)(N-k)$种，取得(白，白，红)的取法有$k(N-k)(N-k-1)$种，故取法共有

$$k(k-1)(k-2)+k(k-1)(N-k)+k(k-1)(N-k)$$
$$+k(N-k)(N-k-1)=k(N-1)(N-2)$$

种. 所以

$$p=\sum_{k=1}^{N}\frac{1}{N}\cdot\frac{k(N-1)(N-2)}{N(N-1)(N-2)}=\frac{1}{N^2}\cdot\frac{N(N+1)}{2}=\frac{N+1}{2N}.$$

例22 在1,2,…,9等9个数字中，有放回地随机取出n个数，求这n个数的乘积能被10除尽的概率.

解 基本事件数有9^n种. 能被10除尽，则这n个数中应含5与偶数. 用逆事件求出，因为不含5的取法有8^n种，不含偶数的有5^n种，既不含5又不含偶数的有4^n种，于是知含5和偶数的取法有$9^n-8^n-5^n+4^n$种. 所以

$$p=\frac{9^n-8^n-5^n+4^n}{9^n}=1-\left(\frac{8}{9}\right)^n-\left(\frac{5}{9}\right)^n+\left(\frac{4}{9}\right)^n.$$

例23 某人将n个准考证随机地发给n个考生，求至少有一个准考证发对的概率.

解 这类问题称为配对问题. 以A_i记第i个准考证发对考生事

件，则由加法公式，至少有一个准考证发对的事件$A=\sum_{i=1}^{n}A_i$，所以

$$P(A)=P(A_1+A_2+\cdots+A_n).$$

因为
$$P(A_i)=\frac{1}{n},\quad \sum_{i=1}^{n}P(A_i)=1,$$

$$P(A_iA_j)=\frac{1}{n}\cdot\frac{1}{n-1}=\frac{1}{n(n-1)}\ (i\neq j),$$

又
$$\sum_{1\leqslant i<j\leqslant n}P(A_iA_j)=\frac{C_n^2}{n(n-1)}=\frac{1}{2!},$$

$$\sum_{1\leqslant i<j<k\leqslant n}P(A_iA_jA_k)=\frac{C_n^3}{n(n-1)(n-2)}=\frac{1}{3!},$$

$$\vdots$$

$$P(A_1A_2\cdots A_n)=C_n^n/n!=1/n!,$$

所以

$$P(A)=1-1/2!+1/3!-\cdots+(-1)^{n-1}/n!\xrightarrow{n\to\infty}1-\mathrm{e}^{-1}.$$

例24 设有n个人排成一排，求：(1) 甲、乙两人之间恰有$r(0\leqslant r\leqslant n-2)$人的概率；(2) 甲、乙两人相邻的概率.

解 n个人排成一排，基本事件数是$n!$.

(1) 甲、乙两人中间有r人的排法数可以这样考虑：设甲排在第i位，则乙排在第$i+r+1$位，所以i只能取$1,2,\cdots,n-r-1$，共$n-r-1$种排法；其余$n-2$个位置是$n-2$人的全排列，有$(n-2)!$种排法；甲、乙位置可以互换，有C_2^1种排法. 由乘法原理，有利事件数为$C_2^1(n-r-1)(n-2)!$，所以

$$p=C_2^1(n-r-1)(n-2)!/n!=2(n-r-1)/[n(n-1)].$$

(2) 甲、乙两人相邻看作占一个位置，而甲、乙位置可以互换，故有$C_2^1(n-1)!$种排法，所以

$$p=C_2^1(n-1)!/n!=2/n.$$

例25 设n个人排成一圈，求：(1) 按顺时针方向，由甲到乙中间相隔r人的概率；(2) 甲、乙两人相邻的概率.

解 n 个人排成一圈,是环排列.

(1) 因为是环排列,就没有首尾之分. 甲、乙按顺时针方向排列,中间相隔 r 人的基本事件数是 n 人取 2 人(甲与乙)的排列,有 P_n^2 种,而甲的位置有 n 种选法,所以

$$p=n/P_n^2=1/(n-1).$$

(2) 设想甲、乙占一个位置,甲、乙可以互换,则甲、乙相邻有 $2(n-2)!$ 种排法. 而一圈只有 $(n-1)!$ 种排法,所以

$$p=2(n-2)!/(n-1)!=2/(n-1).$$

更简单的想法是:一圈有 n 个位置,甲占了一个位置后,乙还有 $n-1$ 个位置可供选择,而与甲相邻的仅两个位置,所以

$$p=2/(n-1).$$

三、几何型概率问题

几何型概率保留了古典型概率的等可能性特征,但样本点的个数为无限(不可列)个. 要根据具体问题选择恰当的几何测度,然后计算事件的概率. 怎样选择几何测度,可详见本节疑难解析 5.

例 26 某轻轨车站每隔 5 min 有一轻轨车通过,乘客随机地来到车站候车,求乘客候车时间不大于 3 min 的概率.

解 由于乘客在 5 min 内的任一时刻到达都是等可能的,符合几何型概率等可能性和无限(不可列)性. 同时,只有一个因素——时间 t 在变,所以用几何量长度来测度. 由题意,得

$$p=\frac{L(A)}{L(\Omega)}=\frac{3}{5}.$$

例 27 将长度为 a 的线段任意分为三段(见图 1.6),求此三线段能构成三角形的概率.

图 1.6

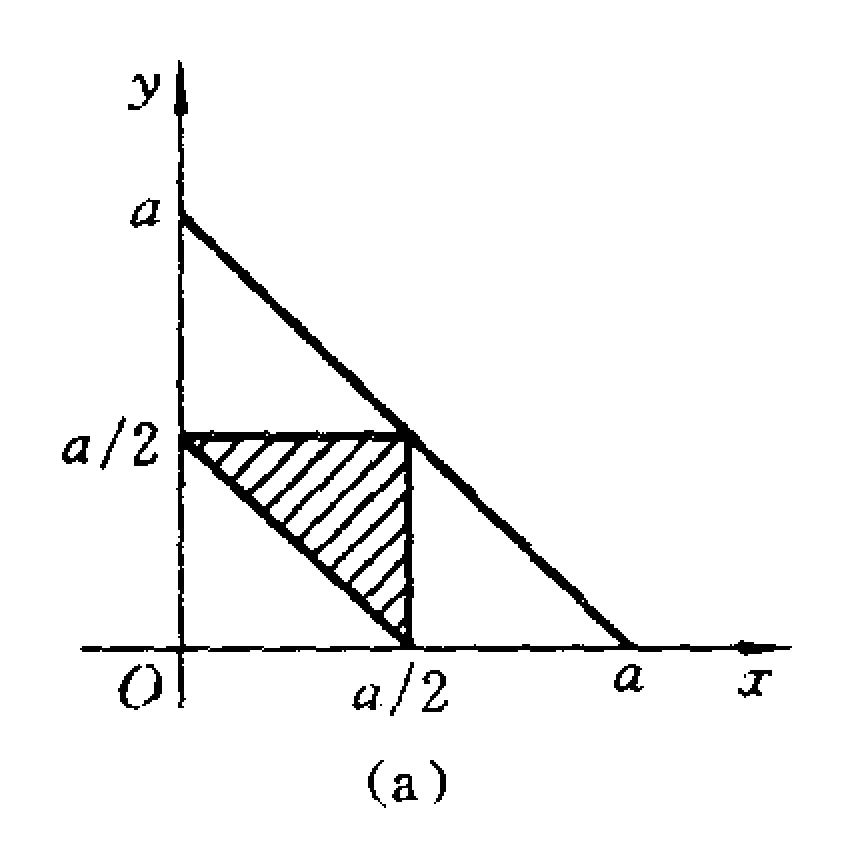
解 设三线段长为 $x,y,a-x-y$,有两个因素 x,y 变化,所以用几何量面积来测度.

(1) 由题意,有 $0<x<a, 0<y<a, 0<x+y<a$,满足此条件的点充满三角形 AOB 内. 而满足构成三角形的点可这样求得:由边的关系,得

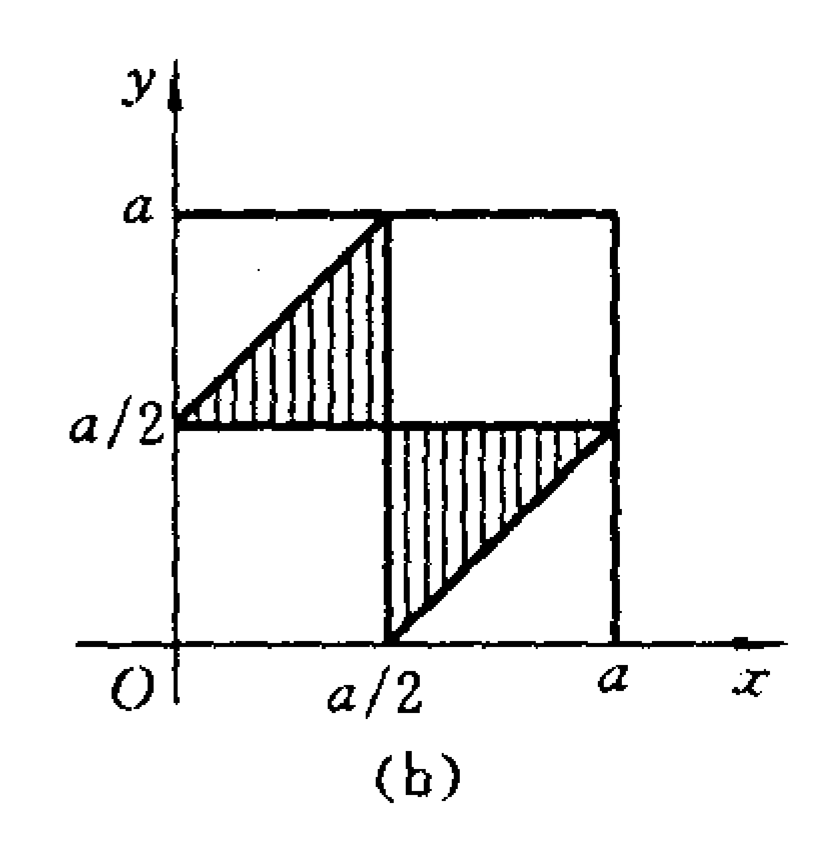
$$\begin{cases}x+y>a-x-y,\\(a-x-y)+y>x,\\(a-x-y)+x>y,\end{cases}\quad 即\quad \begin{cases}x+y>a/2,\\x<a/2,\\y<a/2.\end{cases}$$

满足上述条件的点充满图1.7(a)中阴影域内，故

$$p=\frac{S(A)}{S(\Omega)}=\frac{阴影域的面积}{大三角形的面积}=\frac{1}{4}.$$



(a)



(b)

图1.7

(2) 也可以这样求解：因为 $0<x<a$，$0<y<a$，满足条件的点充满正方形域. 又由组成三角形的边的关系，知 $|y-x|<a/2$，即 $x-a/2<y<x+a/2$. 再考虑 $x<a/2$，$y>a/2$，则组成三角形的点 (x,y) 充满图1.7(b)中左上阴影域；$x>a/2$，$y<a/2$，则组成三角形的点充满右下阴影域. 所以

$$p=\frac{S(A)}{S(\Omega)}=\frac{阴影域的面积}{边长为a的正方形的面积}=\frac{1}{4}.$$

例28(会面问题) 两个朋友约定晚8:00至9:00在某地会面，若先到者等候20 min而另一人不到，先到者则离去，求这对朋友能会面的概率.

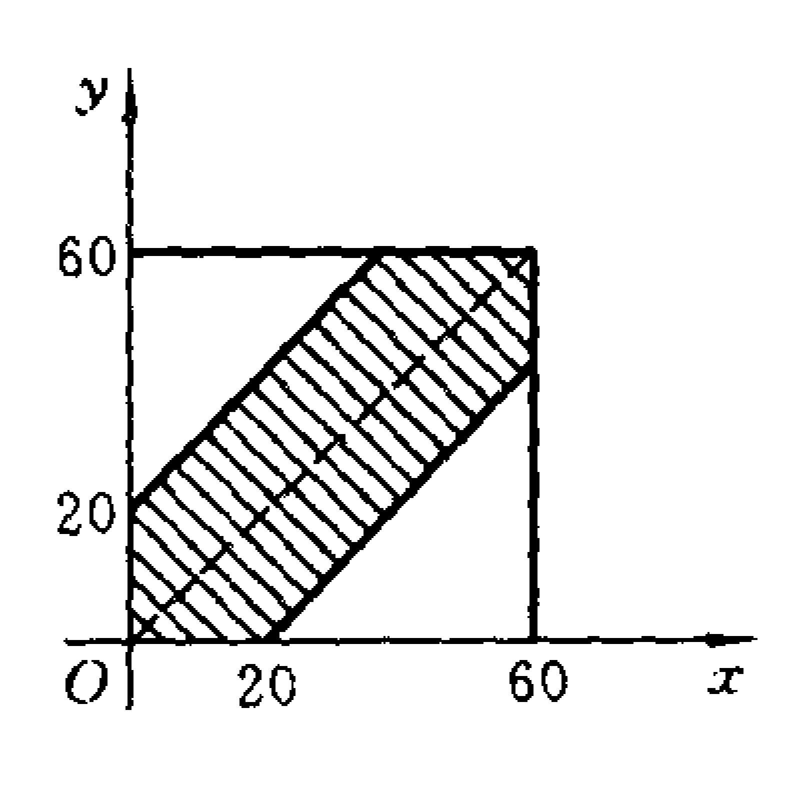


图1.8

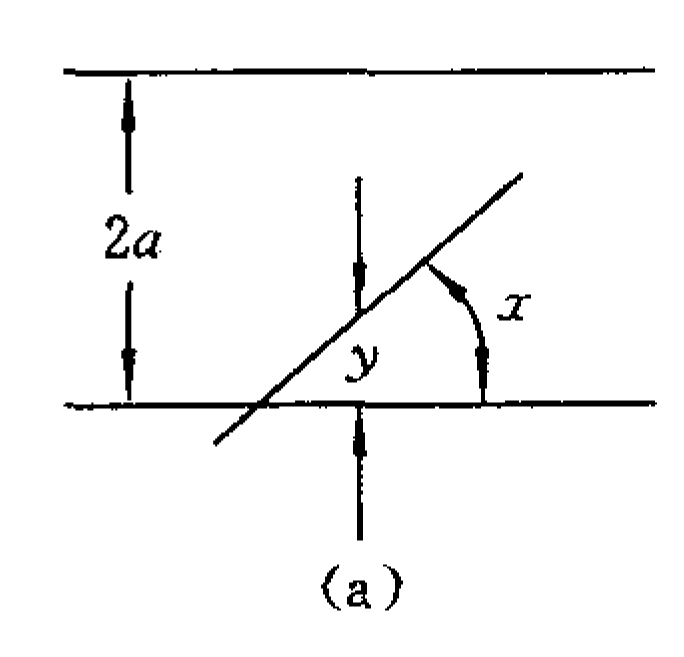
解 这显然是一个几何型概率问题. 以 x,y 表示两人到达时间，则有 $0\leqslant x\leqslant 60$，$0\leqslant y\leqslant 60$，样本点 (x,y) 充满正方形域. 能会面条件 $|x-y|\leqslant 20$，满足条件的点充满图1.8中阴影域. 所以

$$p=\frac{S(A)}{S(\Omega)}=\frac{\text{阴影域的面积}}{\text{正方形的面积}}=\frac{60^2-40^2}{60^2}=\frac{5}{9}.$$

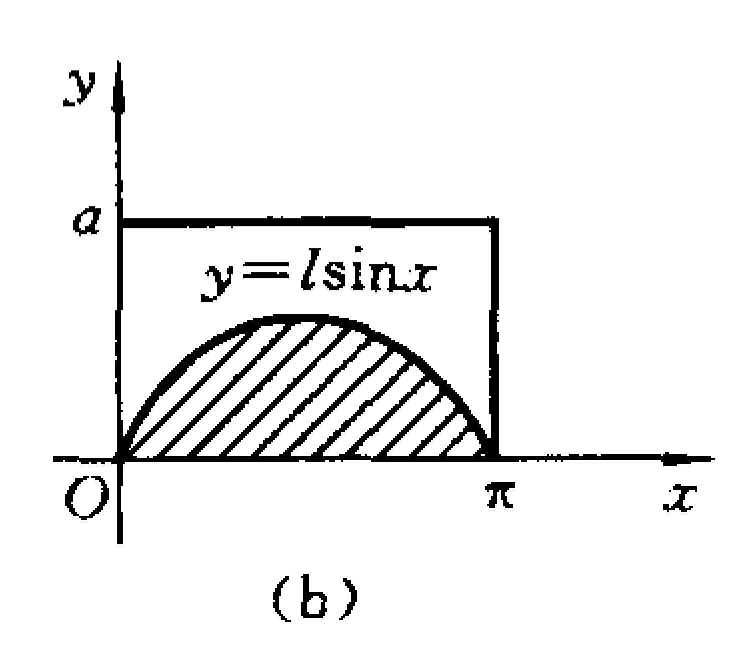
会面问题常用于飞船对接、导弹拦截、码头空出等模型.

例 29(蒲丰投针问题) 设平面上一系列平行线的间距为 $2a$,向平面投一长为 $2l$ 的针($l<a$),求针与平行线相交的概率.

解 如图 1.9(a)所示,以 y 表示针的中点到最近一条平行线的距离,x 表示针与直线的交角,则$0\leqslant y\leqslant a$,$0\leqslant x\leqslant \pi$,样本空间的点$(x,y)$充满图 1.9(b)中的矩形区域.



(a)



(b)

图 1.9

要针与平行线相交,应有 $y\leqslant l\sin x$,满足条件的点充满图 1.9(b)中阴影区域. 所以

$$p=\frac{S(A)}{S(\Omega)}=\frac{\text{阴影域的面积}}{\text{矩形的面积}}=\frac{1}{a\pi}\int_0^{\pi}l\sin x\mathrm{d}x=\frac{2l}{a\pi}.$$

例 30 甲、乙两人相约下午 1:00 到 2:00 之间到某站乘公共汽车外出,他们到车站的时间是随机的. 设在1:00 到2:00 间有四班客车开出,开车时间分别为1:15、1:30、1:45、2:00. 求他们在下述情况下同坐一班车的概率:(1) 约定见车就乘;(2) 约定最多等一班车.

解 设甲、乙到站时间分别为 x,y,则 $1\leqslant x\leqslant 2$,$1\leqslant y\leqslant 2$. 样本点(x,y)充满正方形区域[1,2;1,2],正方形区域可分为 16 个小方格,如图1.10所示.

(1) 见车就乘的情形反映为图上画有交叉线的阴影区域,所以

$$p=4/16=1/4.$$

(2) 约定最多等一班车的情形反映为图上的阴影区域,所以

$$p=10/16=5/8.$$

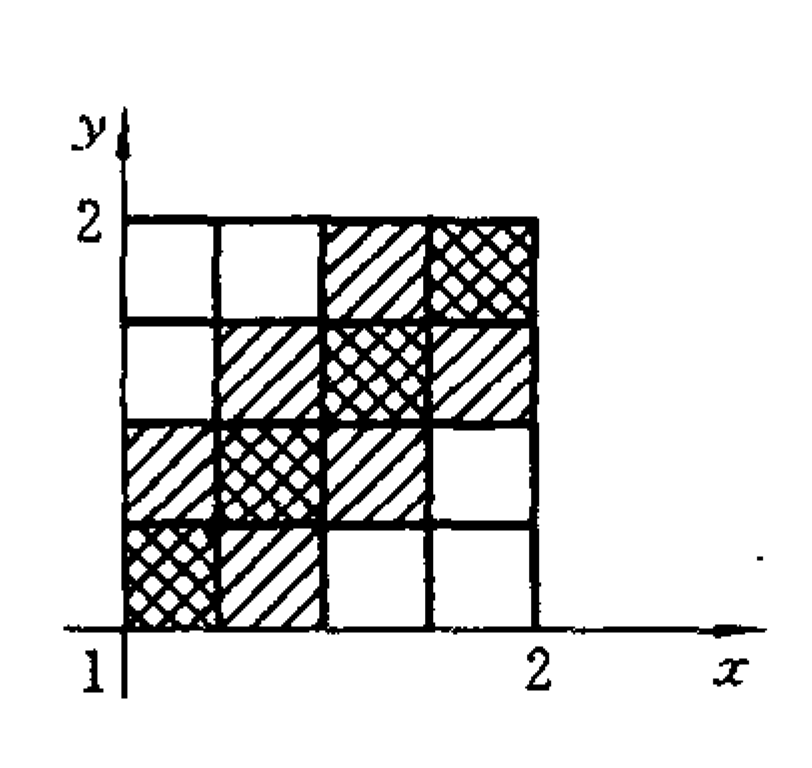


图 1.10

图 1.11

例31 在线段$[0,a]$上任意取三个点x,y,z,求三线段长x,y,z能构成三角形的概率.

解 因为$0\leqslant x\leqslant a,0\leqslant y\leqslant a,0\leqslant z\leqslant a$,所以样本点$(x,y,z)$充满以$a$为边长的长方体(见图1.11). 几何测度取为体积.

x,y,z能构成三角形,则还应满足:

$z<x+y$,即点(x,y,z)在OAC平面下方;

$y<x+z$,即点(x,y,z)在OBC平面上方;

$x<y+z$,即点(x,y,z)在OAB平面上方.

所以点(x,y,z)充满六面体$OABCD$内部.

$$p=\frac{V(A)}{V(\Omega)}=\frac{\text{六面体}OABCD\text{的体积}}{\text{边长为}a\text{的正六面体的体积}}$$
$$=\frac{a^3-3\times 1/3\times a^2/2\times a}{a^3}=\frac{1}{2}.$$

第三节 条件概率与全概率公式

主要内容

1. 条件概率

设A,B是试验E的两个事件,且$P(A)\neq 0$,则在事件B发生的

条件下 A 发生的概率称为事件 A 在条件 B 下的条件概率，记为 $P(A|B)$.

2. 条件概率的计算方法

(1) 在试验 E 的样本空间中计算

$P(A|B)=P(AB)/P(B)$，当 $P(B)>0$ 时；

$P(B|A)=P(AB)/P(A)$，当 $P(A)>0$ 时.

(2) 在缩减样本空间中计算　此时样本空间由 Ω 缩减为 B，有利事件为 AB，所以

$$P(A|B)=P(AB)\ (\text{这时 } P(B)=1).$$

3. 乘法定理

乘法定理是求积事件概率的基本定理.

(1) 若 $P(A)>0$，则 $P(AB)=P(A)P(B|A)$；

(2) 若 $P(AB)>0$，则 $P(ABC)=P(A)P(B|A)P(C|AB)$；

(3) 若 $P(A_1A_2\cdots A_{n-1})>0$，则

$$P(A_1A_2\cdots A_n)$$
$$=P(A_1)P(A_2|A_1)P(A_3|A_1A_2)\cdots P(A_n|A_1A_2\cdots A_{n-1}).$$

4. 完备事件组

设 Ω 为 E 的样本空间，$B_1,B_2,\cdots,B_n$ 为 E 的一组事件，若

(1) $B_iB_j=\varnothing$，$i\neq j$，且 $i,j=1,2,\cdots,n$，

(2) $\sum_{i=1}^{n}B_i=\Omega$，且 $P(B_i)>0$，

则称 $B_1,B_2,\cdots,B_n$ 为 Ω 的一个完备事件组，又称之为 Ω 的一个划分.

5. 全概率公式

设 $B_1,B_2,\cdots,B_n$ 为 Ω 的一个完备事件组，若对 E 的一个事件 A，有 $P(B_i)>0\ (i=1,2,\cdots,n)$，则有全概率公式

$$P(A)=\sum_{i=1}^{n}P(B_i)P(A|B_i).$$

6. 贝叶斯公式

设 $B_1,B_2,\cdots,B_n$ 为 Ω 的一个完备事件组，若对 E 的一个事件

A,有 $P(A)>0$,$P(B_i)>0$ $(i=1,2,\cdots,n)$,则有贝叶斯公式

$$P(B_i|A)=\frac{P(A|B_i)P(B_i)}{\sum_{j=1}^{n}P(B_j)P(A|B_j)},\ i=1,2,\cdots,n.$$

疑难解析

1. 条件概率为什么是概率？它与无条件概率有何区别？

答 条件概率 $P(A|B)$ 可以认为是 B 发生的条件下缩减样本空间 B 中事件 A 的概率,可以验证它符合概率的三条性质:非负性、规范性和可列可加性,因此它是一个概率.

条件概率是在试验 E 的条件下又加上一个新条件(如 B 发生)时,求事件(如 A)的概率. 因为条件增多,则可以理解为:(1) 样本空间 Ω 不变,有利事件改变(由 A 变为 AB,一般地有 $A\supset AB$);(2) 或样本空间 Ω 缩减为 B,有利事件 A 改变为 AB. 所以,一般有

$$P(A)\neq P(A|B).$$

2. 条件概率 $P(A|B)$ 与积事件概率 $P(AB)$ 有何区别与联系？

答 条件概率 $P(A|B)$ 是在试验 E 的条件下增加条件 B 发生后,求得的事件 A 发生的概率. 而积事件 $P(AB)$ 是在试验 E 的条件下 AB 同时发生的概率. 它们的区别就在于,它们发生时的条件不同(虽然都是 A,B 同时发生).

其联系是 $P(A|B)$ 与 $P(AB)$ 可以相互表示,即

$$P(A|B)=P(AB)/P(B),\quad P(AB)=P(B)P(A|B).$$

3. 全概率公式与贝叶斯公式有何联系？它们反映什么样的概率问题？

答 全概率公式与贝叶斯公式是计算复杂事物概率的重要工具.

若把全概率公式中的 A 视为“果”,而把 Ω 的每一划分 B_i 视为

“因”，则全概率公式反映“由因求果”的概率问题. 公式 $P(A)=\sum_{i=1}^{n}P(B_i)P(A|B_i)$ 中的 $P(B_i)$ 是根据以往信息和经验得到的，所以被称为先验概率. 而贝叶斯公式又称为“执果溯因”的概率问题，即在结果 A 已发生的情况下，寻找 A 发生的原因. 公式 $P(B_i|A)=P(A|B_i)P(B_i)/P(A)$ 中的 $P(B_i|A)$ 是得到“信息”A 后求出的，所以被称为后验概率.

先验概率与后验概率有不可分割的联系，后验概率的计算是以先验概率为基础的. 由贝叶斯公式知，求 $P(B_i|A)$ 要用到 $P(A)$，而 $P(A)$ 是由先验概率计算得到的.

方法、技巧与典型例题分析

一、条件概率问题

求解条件概率的关键是能正确地运用条件. 分析条件的发生对样本空间的影响，然后再来考虑计算条件概率，必然可以减少或避免错误.

条件概率的计算方法主要有两种：一是利用公式 $P(A|B)=P(AB)/P(B)$ 计算，要注意到 $P(AB)$ 与 $P(B)$ 都是在样本空间 Ω 中计算；二是在缩减样本空间 B 中计算，这里 $P(B)=1$，$P(A)=P(AB)$. 其它方法还有用定义或直观意义求条件概率的，但不常用. 对于较复杂的条件概率问题，则要善于利用概率的运算性质，化复杂为简单.

例1 已知事件 A，B 互不相容，且 $P(\overline{B})\neq 0$，求概率 $P(A|\overline{B})$.

解 因为 A，B 互不相容，所以 $P(AB)=0$，有

$$P(A|\overline{B})=\frac{P(A\overline{B})}{P(\overline{B})}=\frac{P(A)-P(AB)}{1-P(B)}=\frac{P(A)}{1-P(B)}.$$

例2 n 个人排成一队，已知甲总排在乙的前面，求乙恰好紧跟在甲后面的概率.

解 以 A 记甲排在乙前面事件，以 B 记乙紧跟在甲后面的事件，则要求概率 $P(B|A)$.

在不附带条件的排队中，$P(A)=P(\overline{A})=1/2$. 由于样本空间有$n!$ 个基本事件，A 有$n!/2$ 个基本事件，AB 有$(n-1)!$ 个基本事件(将甲乙看作一个人的全排列，甲、乙不能交换)，则

$$P(B|A)=\frac{P(AB)}{P(A)}=\frac{(n-1)!}{n!}\Big/\frac{n!/2}{n!}=\frac{2}{n}.$$

若在缩减样本空间中计算，则

$$P(B|A)=(n-1)!\Big/\frac{n!}{2}=\frac{2}{n}.$$

例3 电影院的票价为每张5 元，今有$m+n$ 个人排队购票，设有m 人持有5 元币，其余n ($n\leqslant m$)人只有10 元币. 若每人限购一张票，且售票处没有准备零币，求无人等待找钱的概率.

解 以$a_1,a_2,\cdots,a_n$ 记n 个持10 元币的人，以$b_1,b_2,\cdots,b_m$ 记m 个持5 元币的人. 要没有人等待找钱，则在第i 个持10 元币的人之前必须有多于i 个持5 元币的人，所以是条件概率问题.

设 A_k 为 a_k 不等待找钱的事件，则

$$P(A_1)=\frac{m}{m+1}$$

(m 个持5 元币人加一个持10 元币人a_1，a_1 不排在第一位).

$$P(A_2|A_1)=\frac{m-1}{m},$$

$$P(A_3|A_1A_2)=\frac{m-2}{m-1},$$

$$\vdots$$

$$P(A_n|A_1A_2\cdots A_{n-1})=\frac{m-n+1}{m-n+2},$$

所以，没有人等待找钱的概率为

$$\begin{aligned}P(A_1A_2\cdots A_n)&=P(A_1)P(A_2|A_1)\cdots P(A_n|A_1A_2\cdots A_{n-1})\\&=\frac{m}{m+1}\cdot\frac{m-1}{m}\cdot\cdots\cdot\frac{m-n+1}{m-n+2}\end{aligned}$$

$$=\frac{m-n+1}{m+1}.$$

例4 掷三枚骰子，已知得到的三个点数不同，求其中含有点1的概率.

解 用条件概率定义求解. 设 A 为三个点数不同的事件，AB 为其中含有点 1 的事件.

样本空间的基本事件数为 $6^3=216$，A 包含基本事件数为 $P_6^3=120$，AB 含基本事件数 $P_3^3C_5^2=60$，所以

$$P(A)=120/216=5/9,\quad P(AB)=60/216=5/8,$$

$$P(B|A)=60/120=1/2.$$

例5 一袋中有 r 个红球，t 个白球，每次从袋中任取一个球观察后放回，同时放入 a 个与其同色的球. 在袋中连续取球四次，求第一、二次取到红球而第三、四次取到白球的概率.

解 以 A_i $(i=1,2,3,4)$记第 i 次取到红球的事件，则所求事件为 $A_1A_2\overline{A}_3\overline{A}_4$，其概率为

$$\begin{aligned}P(A_1A_2\overline{A}_3\overline{A}_4)&=P(A_1)P(A_2|A_1)P(\overline{A}_3|A_1A_2)P(\overline{A}_4|A_1A_2\overline{A}_3)\\&=\frac{r}{r+t}\cdot\frac{r+a}{r+t+a}\cdot\frac{t}{r+t+2a}\cdot\frac{t+a}{r+t+3a}\\&=\frac{rt(r+a)(t+a)}{(r+t)(r+t+a)(r+t+2a)(r+t+3a)}.\end{aligned}$$

例6 已知某家庭有3个孩子，其中至少有1个是女孩. 求这个家庭至少有1个男孩的概率(设生男生女的概率相等).

解 以 A 记至少有 1 个女孩事件，以 B 记至少有 1 个男孩事件，求 $P(B|A)$用公式 $P(B|A)=P(AB)/P(A)$.

由 $P(\overline{A})=(1/2)^3=1/8$，得 $P(A)=7/8$，$P(B)=7/8$.

由 $\overline{A}$ 与 $\overline{B}$ 互不相容，$\overline{AB}=\overline{A}+\overline{B}$，得

$$P(\overline{AB})=P(\overline{A})+P(\overline{B})=1/4\Longrightarrow P(AB)=3/4,$$

所以 $$P(B|A)=P(AB)/P(A)=(3/4)/(7/8)=6/7.$$

例7 已知n 件产品中有m 件次品，从中任取2 件. 在得知其中1 件是次品的情况下，求另一件也是次品的概率.

解 以 A_1 记 2 件中至少有 1 件次品的事件，以 A_2 记 2 件全是次品的事件，要求 $P(A_2|A_1)$.

因为 $A_2 \supset A_1$，所以

$$A_1A_2 = A_2, \quad P(A_1A_2) = P(A_2).$$

而
$$P(A_1) = 1 - P(\overline{A}_1) = 1 - C_{n-m}^2 / C_n^2$$
$$= [m(2n-m-1)]/[n(n-1)],$$
$$P(A_2) = C_m^2 / C_n^2 = [m(m-1)]/[n(n-1)],$$

所以
$$P(A_2|A_1) = P(A_1A_2)/P(A_1) = P(A_2)/P(A_1)$$
$$= (m-1)/(2n-m-1).$$

例8 已知掷 5 枚硬币时至少出现 2 个正面，求正面数恰为 3 个的概率.

解 以 B 记至少出现2 个正面事件，以 A_i 记恰好出现 i 个正面事件，则 $B = A_2 + A_3 + A_4 + A_5$.

由于 A_i 互不相容，$B \supset A_3$，$A_3B = A_3$，而

$$P(A_3) = C_5^3/2^5, \quad P(B) = (C_5^2 + C_5^3 + C_5^4 + C_5^5)/2^5$$

(因为掷 5 枚硬币基本事件为 2^5 个，A_i 含基本事件 C_5^i 个)，所以

$$P(A_3|B) = P(A_3B)/P(B) = P(A_3)/P(B)$$
$$= C_5^3/(C_5^2 + C_5^3 + C_5^4 + C_5^5) = 5/13.$$

二、全概率公式与贝叶斯公式问题

在求解全概率公式与贝叶斯公式问题时，表面上似乎只要套公式就可以了，其实很容易出错. 因此，特别要注意以下两点：(1) 寻找一个正确的样本空间的完备事件组(划分) $B_1, B_2, \cdots, B_n$，并准确计算 $P(B_i)$ 和 $P(A|B_i)$ 的值；(2) 不要混淆 $P(A|B_i)$ 与 $P(B_i|A)$，因为两者完全不同.

例9 甲、乙、丙三个工厂生产了一批同样规格的零件，它们的产量分别占总产量的 20%，40%，40%，它们的次品率分别为 5%，4%，3%. 今从仓库中任取 1 个零件，它是次品的概率为多少?

解 以 A_1, A_2, A_3 记分别取到甲、乙、丙厂生产的零件的事件，以 B 记零件为次品的事件，则

$$P(A_1)=0.2,\quad P(A_2)=0.4,\quad P(A_3)=0.4,$$

$$P(B|A_1)=0.05,\quad P(B|A_2)=0.04,\quad P(B|A_3)=0.03.$$

因为 A_1,A_2,A_3 为一个完备事件组，由全概率公式，有

$$P(B)=\sum_{i=1}^{3}P(A_i)P(B|A_i)$$
$$=0.2\times0.05+0.4\times0.04+0.4\times0.03=0.038.$$

例 10 有两箱同型号的零件，A 箱内装 50 件，其中一等品 10 件；B 箱内装 30 件，其中一等品 18 件. 装配工从两箱中任选一箱，从箱子中先后随机地取两个零件（不放回抽样）. 求：

(1) 先取出的一件是一等品的概率；

(2) 在先取出的一件是一等品的条件下，第二次取出的零件仍是一等品的概率.

解 以 A,B 分别记从 A 箱、B 箱内取得一等品的事件. 由已知条件 $P(A)=P(B)=1/2$，并设 C_i 为第 i $(i=1,2)$ 次取得一等品的事件.

(1) 因为 $P(C_1|A)=1/5, P(C_1|B)=3/5$，依全概率公式，得

$$P(C_1)=P(A)P(C_1|A)+P(B)P(C_1|B)$$
$$=\frac{1}{2}\left(\frac{1}{5}+\frac{3}{5}\right)=\frac{2}{5}.$$

(2) 由条件概率公式与全概率公式

$$P(C_2|C_1)=P(C_1C_2)/P(C_1)$$
$$=[P(A)P(C_1C_2|A)+P(B)P(C_1C_2|B)]/P(C_1)$$
$$=\frac{1}{2}\left(\frac{10\times9}{50\times49}+\frac{18\times17}{30\times29}\right)\Big/\frac{2}{5}=0.4856.$$

例 11 已知某批产品的合格率为 0.9. 检验员检验时，将合格品误认为次品的概率为 0.02，而一个次品被误认为合格的概率为 0.05. 求：

(1) 检查任一产品被认为是合格品的概率；

(2) 被认为合格品的产品确实合格的概率.

解 以 B 记一个产品检查被认为合格的事件，以 A 记产品确

实合格的事件，则 $A,\overline{A}$ 构成一个完备事件组，$P(A)=0.9,P(\overline{A})=0.1,P(B|A)=0.98,P(B|\overline{A})=0.05$. 于是：

(1) 由全概率公式，一个产品被认为合格的概率为

$$P(B)=P(A)P(B|A)+P(\overline{A})P(B|\overline{A})$$
$$=0.9\times0.98+0.1\times0.05=0.887.$$

(2) 由贝叶斯公式，被认为合格的产品确实合格的概率为

$$P(A|B)=[P(A)P(B|A)]/P(B)=0.9\times0.98/0.887$$
$$=0.994.$$

例 12 某人到武汉参加会议，他乘火车、轮船、汽车或飞机去的概率分别为 0.2，0.1，0.3 和 0.4. 如果他乘火车、轮船、汽车前去，迟到的概率分别为 1/3，1/12 和 1/4，乘飞机不会迟到. 结果他迟到了，求他乘汽车去的概率.

解 以 B 记开会迟到事件，以 A_1,A_2,A_3,A_4 分别记某人乘火车、轮船、汽车和飞机去的事件，则 A_1,A_2,A_3,A_4 为一完备事件组. 由全概率公式

$$P(B)=\sum_{i=1}^{4}P(A_i)P(B|A_i)$$
$$=0.2\times1/3+0.1\times1/12+0.3\times1/4+0.4\times0$$
$$=0.15.$$

又由贝叶斯公式，迟到是因为乘汽车的概率是

$$P(A_3|B)=(0.3\times1/4)/0.15=0.5.$$

例 13 某炮兵阵地有甲、乙、丙三门炮，三门炮的命中率分别为 0.4，0.3，0.5，设三门炮同时向一目标发射炮弹，结果共有两发击中该目标，求此时是甲炮击中的概率.

解 这是一个“执果溯因”问题. 以 A_1,A_2,A_3 分别记甲、乙、丙击中的事件，则 $P(A_1)=0.4,P(A_2)=0.3,P(A_3)=0.5$. 又以 A 记三发炮弹有两发击中的事件，以 B 记三炮齐射甲炮击中的事件，则所求为 $P(B|A)$.

若事件 B 发生，则另一发命中的为乙炮或丙炮所发炮弹，故

$$P(A|B)=P(A_2)[1-P(A_3)]+[1-P(A_2)]P(A_3)$$
$$=0.3\times0.5+0.7\times0.5=0.5,$$
$$P(A|\overline{B})=P(A_2)P(A_3)=0.3\times0.5=0.15,$$

所以

$$P(B|A)=[P(B)P(A|B)]/[P(B)P(A|B)+P(\overline{B})P(A|\overline{B})]$$
$$=(0.4\times0.5)/(0.4\times0.5+0.6\times0.15)=20/29.$$

例 14 设有 24 个外形相同的球分装在三个盒子内，每盒装 8 个. 第一盒内有 5 个标有 A 的球，3 个标有 B 的球；第二盒内有红球和白球各 4 个；第三盒内有红球 6 个，白球 2 个. 先从第一盒中取 1 个球，若是 A 字球，则在第二盒中任取 1 个球；若是 B 字球，则从第三盒中任取 1 个球，求第二次取得的球是红球的概率.

解 以 A 记第一次从第一盒中取得 A 字球事件，以 B 记第一次从第一盒中取得 B 字球事件，以 C 记第二次取出红球事件，则可利用“概率树”图形（见图 1.12）分析表示事件的关系.

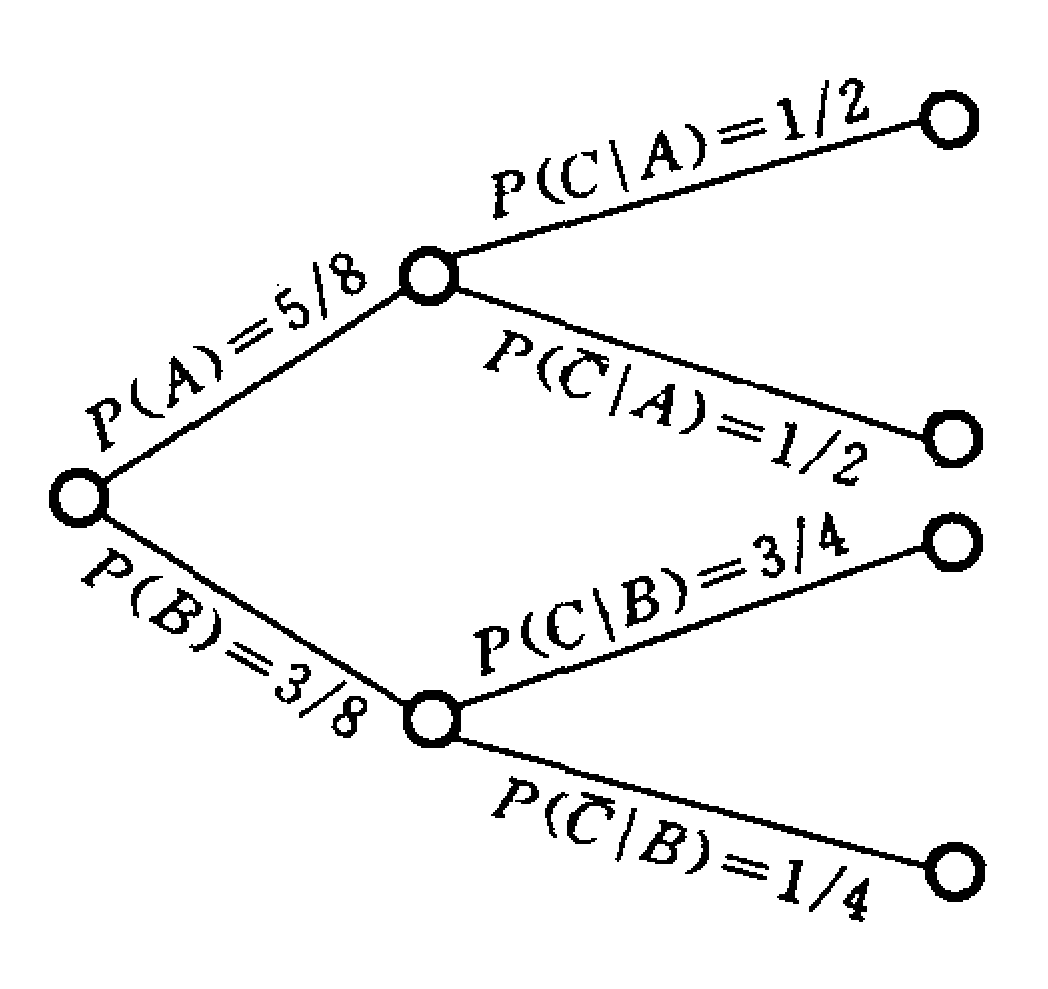


图 1.12

依题意有

$$P(A)=5/8,\quad P(B)=3/8;$$
$$P(C|A)=1/2,\quad P(\overline{C}|A)=1/2;$$
$$P(C|B)=3/4,\quad P(\overline{C}|B)=1/4.$$

由全概率公式，得

$$P(C)=P(C|A)P(A)+P(C|B)P(B)$$
$$=1/2\times5/8+3/4\times3/8=19/32.$$

概率树是用来分析复杂事件中事物间关系的一种既形象又有效的方法，凡是利用全概率公式求解的问题，都可以借助概率树直观、清晰地表示概率关系.

例 15 某卫生机构的资料表明：患肺癌的人中吸烟的占 90%，不患肺癌的人中吸烟的占 20%. 设患肺癌的人占人群的 0.1%，求在吸烟的人中患肺癌的概率.

解　以A记被观察者吸烟的事件，以B_1记被观察者中患肺癌的事件，以B_2记被观察者中不患肺癌的事件，则B_1,B_2为一完备事件组.依题意有，$P(B_1)=0.001$，$P(B_2)=0.999$，$P(A|B_1)=0.9$，$P(A|B_2)=0.2$，所以，由贝叶斯公式，得

$$
\begin{aligned}
P(B_1|A)&=\frac{P(B_1)P(A|B_1)}{P(B_1)P(A|B_1)+P(B_2)P(A|B_2)}\\
&=(0.001\times0.9)/(0.001\times0.9+0.999\times0.2)\\
&=0.0045.
\end{aligned}
$$

显然，患肺癌的人中吸烟者的比例很高不等于吸烟者中患肺癌的比例很高，即$P(B_1|A)\neq P(A|B_1)$，故不能以患肺癌的人中吸烟者比例很高而认为吸烟者患肺癌的比例很高.

第四节　独立性与伯努利概型

主 要 内 容

一、独立性

1. 两事件相互独立

对随机试验E的两个事件A,B，若$P(AB)=P(A)P(B)$，则称A,B为相互独立的事件，也简称事件A,B独立.

(1) 若A,B相互独立，则A与$\overline{B}$，$\overline{A}$与B，$\overline{A}$与$\overline{B}$也相互独立.

(2) 若A,B相互独立，则$P(B|A)=P(B)$，$P(A|B)=P(A)$.反之亦然.

2. 三事件相互独立

若对A,B,C三事件，以下等式成立：

$$P(AB)=P(A)P(B),\quad P(BC)=P(B)P(C),$$

$$P(AC)=P(A)P(C),\quad P(ABC)=P(A)P(B)P(C),$$

则称 A,B,C 相互独立.

若仅前三个等式成立,最后一个等式不成立,则称 A,B,C 两两独立.

3. n 个事件相互独立

若对 n 个事件 $A_1,A_2,\cdots,A_n$,对于任意的 k $(1<k\leqslant n)$ 及 $1\leqslant i_1<i_2<\cdots<i_k\leqslant n$,等式

$$P(A_{i_1}A_{i_2}\cdots A_{i_k})=P(A_{i_1})P(A_{i_2})\cdots P(A_{i_k})$$

成立,则称 $A_1,A_2,\cdots,A_n$ 相互独立.

上面的等式实际含有 2^n-n-1 个等式.

二、伯努利概型

若将试验 E 重复进行 n 次,每次试验的结果互不影响,则称 n 次试验是相互独立的.

若试验 E 只有两个可能结果,A 或 $\overline{A}$,则称 E 为伯努利试验.将试验 E 独立地重复进行 n 次,则称 E 为 n 重伯努利试验.

在 n 重伯努利试验中,若一次试验时事件 A 发生的概率为 p,则在 n 重伯努利试验中 A 发生 k 次的概率为

$$P\{X=k\}=C_n^kp^k(1-p)^{n-k}.$$

由于 $P\{X=k\}=C_n^kp^k(1-p)^{n-k}$ 恰好是二项式 $(p+q)^n$ 的展开式中含 p^k 的项,所以又称之为二项概率.

疑 难 解 析

1. 两事件 A,B 相互独立与两事件 A,B 互斥这两个概念有什么关系?

答 两者没有必然的联系.两事件 A,B 相互独立,则 A 发生与 B 发生无关,两者互不影响;两事件 A,B 互斥,则 A 发生一定有 B 不发生,两者是相互影响的.

图 1.13 为 A,B 相互独立(图(a))与 A,B 互斥(图(b))的一个例子.可以看出:A,B 相互独立时 AB 可能非空,而 A,B 互斥时一

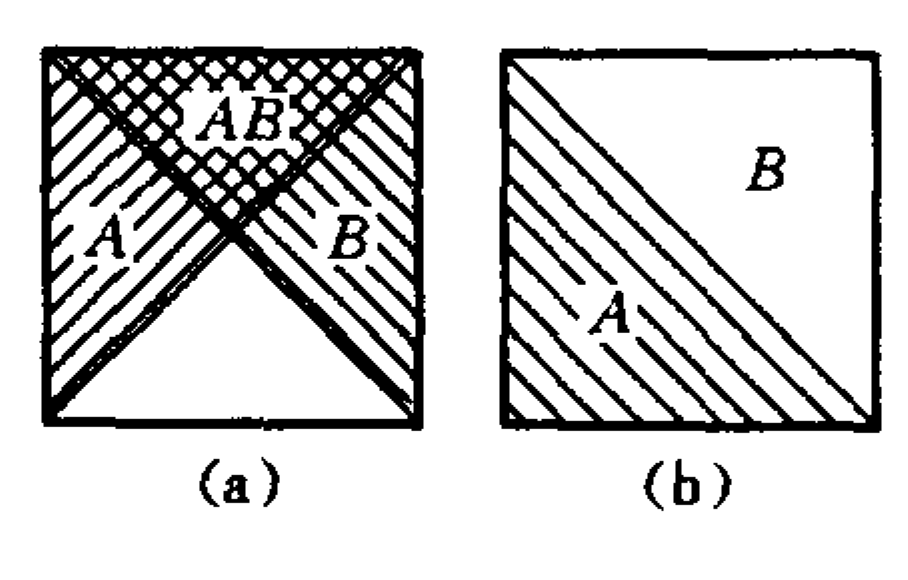


图1.13

定有 $AB=\varnothing$.

2. 有放回抽样和无放回抽样与独立性有什么关系?

答 有放回抽样时,对抽出的样本观察后仍放回,因此样本空间没有发生变化,两次抽样可以看作独立的重复抽样,计算事件的概率时就可以利用独立性概念.无放回抽样时,抽出的样本不再放回,因此前后抽样的样本空间是不同的,不能作为独立的重复抽样处理,计算概率时不能利用独立性概念,一般依据条件概率、乘法公式与其它运算法则处理.

3. 多个事件的两两独立与相互独立有什么不同?

答 对两个事件而言,两两独立与相互独立是一致的,而两个事件以上时,两者就不同了.两两独立是指其中任意两个事件相互独立,而全部事件相互独立是指其中任意个事件都相互独立.如事件组 $A_1,A_2,\cdots,A_n$ 相互独立包括两两独立、三三独立……n 个事件独立.

由 n 个事件两两独立不能得出 n 个事件相互独立,但后者可以推出前者.如掷一个正四面体,其中三面分别写有 A,B,C 字样,另一面写有 ABC 字样.观察朝下一面上的字母,得 $P(A)=P(B)=P(C)=1/2$, $P(AB)=P(A)P(B)=1/4$, $P(AC)=P(A)P(C)=1/4$, $P(BC)=P(B)P(C)=1/4$,但 $P(ABC)=1/4\neq P(A)P(B)P(C)$.足以见得,三事件的两两独立,不能推出三事件的相互独立.

4. 随机事件的相互独立与随机试验的相互独立怎样区分?

答 随机事件的相互独立是指同一随机试验中各个随机事件的发生互不影响、相互独立的性质.随机试验的独立性有两种类型:一是重复试验的独立性,如某人昨天投篮试验的结果与今天投篮试验的结果可以认为是独立的,n 重伯努利试验是随机试验独立的典型例子;二是不同随机试验的独立性,即从不同的随机试验中各取任一随机事件,则它们是相互独立的,如某车间各机床是否

正常工作是独立的，某电路各元件是否可靠也是独立的.

5. 怎样区分试验是否为伯努利试验?

答 只有两个结果的独立重复试验称为伯努利试验. 要注意的是两个结果不等于两个样本点，如射击，成绩可以是 0 环、1 环……10 环，但我们可以分为命中、不中两个结果. 如果只考虑事件 A 是否发生，则可以把试验 E 的其它事件都看作 $\overline{A}$，那么试验 E 就是一个伯努利试验，因此伯努利试验是广泛存在的. 而 n 重伯努利试验是在相同条件下重复进行的伯努利试验，概率的统计定义就是由此概率模型得出的，以后将要讲到的许多分布也与其有关.

方法、技巧与典型例题分析

一、独立性问题

独立性是概率计算的一个重要条件，许多定理与公式仅在独立性条件下才成立. 因此首要的问题是判别所要计算的概率中的各事件是否相互独立. 判别的方法是:(1) 用定义验证，但当事件较多时会比较困难;(2) 根据实际问题，由经验确定是否相互独立. 在实际计算时，要牢记一些常用的公式与结果(见主要内容)，并能熟练运用它们.

例1 加工某一产品有三道工序. 设第一、第二、第三道工序的次品率分别为 2%，3%，5%，假定各道工序相互独立，求完成的产品的次品率.

解 以 A_1，A_2，A_3 记第一、第二、第三道工序合格事件，记 A 为产品为合格品的事件，则 $P(A_1)=0.98$，$P(A_2)=0.97$，$P(A_3)=0.95$. 由乘法公式和独立性，得

$$P(A)=P(A_1A_2A_3)=P(A_1)P(A_2)P(A_3)=0.90307,$$

由逆事件概率，得次品率 $p=1-P(A)=0.09693$.

例 2 某工人在车间照管两台不同的机器，由以往经验知，在某段时间内，甲机器的停工率为 0.15，乙机器的停工率为 0.20. 求

该段时间内至少有一台机器不停工的概率.

解 以 A,B 分别记甲、乙机器不停工的事件,可由经验断定甲、乙机器停工是独立的.所以,所求概率为 $P(A+B)$,且

$$\begin{aligned}P(A+B)&=P(A)+P(B)-P(AB)\\&=P(A)+P(B)-P(A)P(B)\\&=0.85+0.8-0.85\times0.8=0.97.\end{aligned}$$

例3 一门炮对同一目标进行了三次独立的射击,三次射击的命中率分别为0.4,0.5,0.7,求:

(1) 三次射击中恰有一次击中目标的概率;

(2) 三次射击至少有一次击中目标的概率.

解 以 A_i 记第 i 次击中目标的事件,则

$$P(A_1)=0.4,\quad P(A_2)=0.5,\quad P(A_3)=0.7.$$

(1) 以 B 记恰有一次击中目标的事件,得

$$\begin{aligned}P(B)&=P(A_1\overline{A}_2\overline{A}_3\cup\overline{A}_1A_2\overline{A}_3\cup\overline{A}_1\overline{A}_2A_3)\ (\text{由互不相容性})\\&=P(A_1\overline{A}_2\overline{A}_3)+P(\overline{A}_1A_2\overline{A}_3)+P(\overline{A}_1\overline{A}_2A_3)\ (\text{由独立性})\\&=0.4\times0.5\times0.3+0.6\times0.5\times0.3+0.6\times0.5\times0.7\\&=0.36.\end{aligned}$$

(2) 以 C 记至少有一次击中目标的事件,得

$$\begin{aligned}P(C)&=1-P(\overline{C})=1-P(\overline{A}_1\overline{A}_2\overline{A}_3)=1-P(\overline{A}_1)P(\overline{A}_2)P(\overline{A}_3)\\&=1-0.6\times0.5\times0.3=0.91.\end{aligned}$$

例4 某系统如图1.14所示,继电器触点1,2,3,4,5,6闭合的概率均为 p,且各继电器触点的闭合是相互独立的,求系统是通路的概率.

解 将系统分为子系统,子系统又分并联与串联,实行分步骤方法来求系统是通路的概率.

(1) 在第一个子系统中,2与3是串联,线路接通是交事件,由独立性知,接通的概率为 p^2;而1与2,3是并联,接通的概率用逆概率求,为 $1-(1-p)(1-p^2)=p+p^2-p^3$.

第二个子系统接通的概率为 p,与第一个子系统串联,是交事

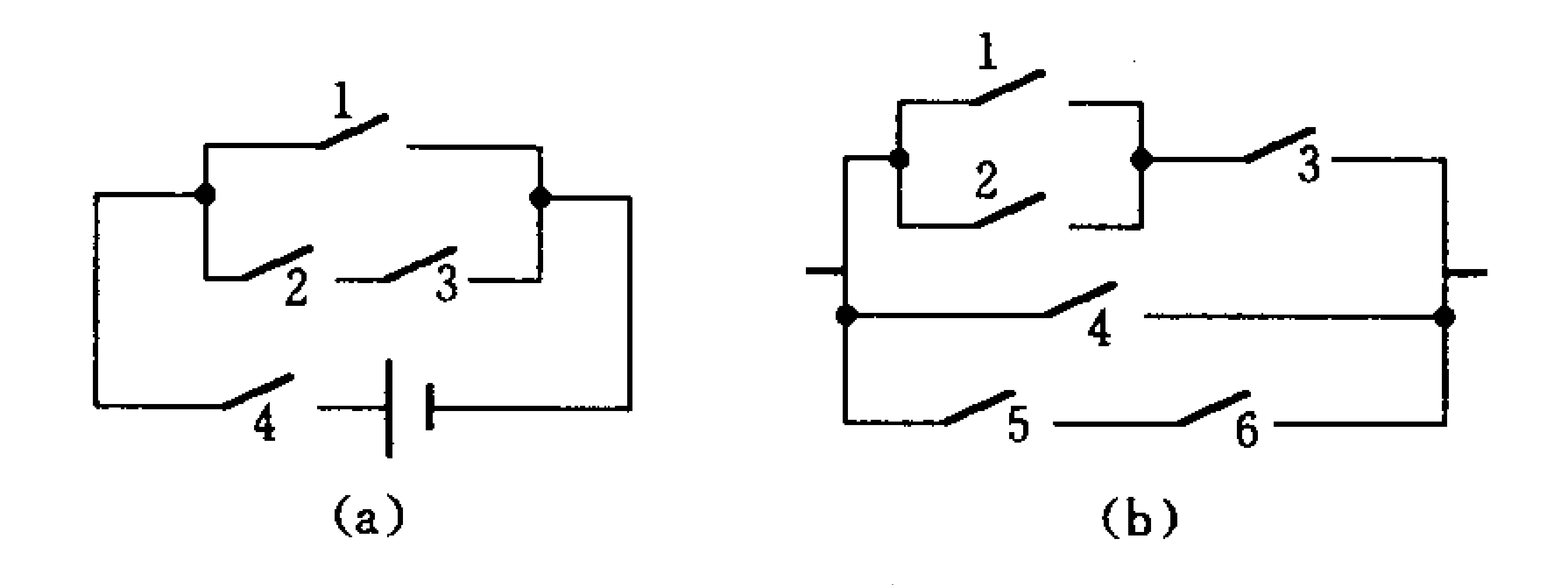


图 1.14

件，故整个系统是通路的概率为

$$p_1=p(p+p^2-p^3)=p^2+p^3-p^4.$$

(2) 系统由三个子系统并联而成，逐个求出子系统线路接通的概率，再合成求整个系统是通路的概率.

在第一个子系统中，1 与 2 是并联，1，2 与 3 是串联，所以，线路接通的概率为 $[1-(1-p)(1-p)]p$；第二个子系统线路接通的概率为 p；第三个子系统由 5，6 串联而成，线路接通的概率为 p^2. 于是，系统是通路的概率为

$$\begin{aligned}p_2&=1-[1-p^2(2-p)](1-p)(1-p^2)\\&=p+3p^2-4p^3-p^4+3p^5-p^6.\end{aligned}$$

例 5 排球比赛的规则是 5 局 3 胜制. A、B 两队的胜率分别为 0.6 和 0.4. 前两局 A 队以 2 : 0 领先，求 A 队获胜的概率.

解 以 A_i $(i=3,4,5)$ 记 A 在第 i 局获胜的事件，以 B 记 A 队获胜的事件，则 $P(A_i)=0.6$. 所以

$$\begin{aligned}P(B)&=P(A_3\cup\overline{A}_3A_4\cup\overline{A}_3\overline{A}_4A_5)\\&=P(A_3)+P(\overline{A}_3)P(A_4)+P(\overline{A}_3)P(\overline{A}_4)P(A_5)\\&=0.6+0.6\times0.4+0.6\times0.4^2=0.936.\end{aligned}$$

例 6 在一批 N 个产品中有 M 个次品，每次任取一个，观察后放回，求：

(1) n 次都取得正品的概率；

(2) n 次中至少有一次取得次品的概率.

解 因为是放回抽样，可以认为各次抽取相互独立. 以 A_i 记

第 i 次取得合格品的事件，则 $P(A_i)=1-M/N$.

(1) $p_1=P(A_1A_2\cdots A_n)=P(A_1)P(A_2)\cdots P(A_n)$

$$=(1-M/N)^n.$$

(2) $p_2=1-p_1=1-(1-M/N)^n$.

例7 设 $P(A)>0, P(B)>0$，证明：

$$A,B\text{ 相互独立}\Longleftrightarrow P(A|B)=P(A|\overline{B}).$$

证 必要性 若 A,B 相互独立，则

$$P(A|B)=A, P(A|\overline{B})=A,$$

所以

$$P(A|B)=P(A|\overline{B}).$$

充分性 若 $P(A|B)=P(A|\overline{B})$，即

$$\frac{P(AB)}{P(B)}=\frac{P(A\overline{B})}{P(\overline{B})}\Longrightarrow\frac{P(AB)}{P(B)}=\frac{P(A)-P(AB)}{1-P(B)},$$

得 $P(AB)=P(A)P(B)$，即 A,B 相互独立.

例8 设 A,B 相互独立，且 $P(A\overline{B})=P(\overline{A}B)=1/4$，求概率 $P(A)$ 与 $P(B)$.

解 由 A,B 相互独立知：$A,\overline{B}$ 相互独立，$\overline{A},B$ 相互独立.

$$P(A\overline{B})=P(A)(1-P(B))=1/4,$$

$$P(\overline{A}B)=P(B)(1-P(A))=1/4,$$

于是

$$P(A)-P(A)P(B)=P(B)-P(B)P(A)$$

$$\Longrightarrow P(A)=P(B),$$

从而

$$P(A)-P(A)^2=1/4\Longrightarrow P(A)=P(B)=1/2.$$

二、伯努利概型问题

在求解伯努利概型问题之前，首先要确认试验是否是 n 重独立试验，再确认试验结果是否只有两个.要求：(1) 确定重数 n 及一次试验中 A 发生的概率 p；(2) 确定所求 A 发生的次数 k(或 $P(A)$的取值范围)；(3) 用二项概率公式计算 $p_n(k)$(或确定 k 值).

例9 求 n 重伯努利试验中 A 成功奇数次的概率.

解 设一次试验中 A 成功的概率为 p，则成功奇数次的概率可利用二项展开式求得.将

$$(q+p)^n=C_n^0p^0q^n+C_n^1pq^{n-1}+C_n^2p^2q^{n-2}+\cdots=1,$$
$$(q-p)^n=C_n^0p^0q^n-C_n^1pq^{n-1}+C_n^2p^2q^{n-2}+\cdots$$

相减，得所求概率为

$$\begin{aligned}p&=C_n^1pq^{n-1}+C_n^3p^3q^{n-3}+\cdots\\&=[1-(q-p)^n]/2=[1-(1-2p)^n]/2.\end{aligned}$$

例10 某商场各柜台受到消费者投诉的事件数为0,1,2三种情形，其概率分别为0.6,0.3,0.1. 有关部门每月抽查商场的两个柜台，规定：如果两个柜台受到投诉之和超过1，则给商场通报批评；若一年中有两个月受到通报批评，则该商场受挂牌处分一年. 求该商场受处分的概率.

解 以A记商场某月受通报批评事件，以B_i $(i=0,1,2)$记第一个柜台受i次投诉的事件，以C_i记第二个柜台受i次投诉的事件，则

$$\begin{aligned}P(A)&=P(B_2C_0+B_0C_2+\overline{B}_0\overline{C}_0)\ (\text{由加法公式、独立性})\\&=P(B_2)P(C_0)+P(B_0)P(C_2)+P(\overline{B}_0)P(\overline{C}_0)\\&=0.1\times0.6+0.6\times0.1+0.4\times0.4=0.28.\end{aligned}$$

以X记一年中受通报批评次数，则

$$\begin{aligned}P\{X\geqslant3\}&=1-P\{X=0\}-P\{X=1\}-P\{X=2\}\\&=1-C_{12}^0\times0.28^0\times0.72^{12}-C_{12}^1\times0.28\times0.72^{11}\\&\quad-C_{12}^2\times0.28^2\times0.72^{10}\\&=0.696.\end{aligned}$$

例11 一批产品数量很大，其次品率为0.05，从中抽出50个进行检验，若次品数超过1个，则认为这批产品是不合格的. 求这批产品被认为是合格的概率.

解 由于产品数量很大，可把一次抽50个视同于有放回地抽50个，当作50重伯努利试验. 因此，这批产品合格，相当于50个产品中次品不超过1个. 由$p=0.05,n=50$，得

$$p=C_{50}^0\times0.05^0\times0.95^{50}+C_{50}^1\times0.05\times0.95^{49}=0.2794.$$

例12 某厂生产了一批产品有15000件，其中次品150件. 现

从产品中无放回地随机抽取100件，求恰有2件次品的概率.

解 这是一个超几何分布的概率模型，$N=15000$，$M=150$，$n=100$，所以

$$p=C_{150}^{2}C_{14850}^{98}/C_{15000}^{100},$$

显然很难计算. 因为n很大，$p=0.01$很小，可用二项分布近似，即

$$p_k\approx C_{100}^{2}\times 0.01^2\times 0.99^{98}$$

仍然不易计算. 但n很大，令$\lambda=np=1$，则可用泊松分布近似(下章将详细讨论)，得

$$p_k\approx 1^2\times e^{-1}/2!.$$

例13 设在每次试验中，事件A发生的概率是p. 进行了n次独立重复试验，问：事件A至少发生一次的概率是多少？若要使n次独立重复试验中A至少发生一次的概率不小于p_1，问：n应取多大？

解 以A_k $(k=0,1,\cdots)$记事件A出现的次数，则A发生一次的概率可用二项概率求得，即

$$p_k=1-C_n^0p^0(1-p^n)=1-(1-p)^n.$$

要使$p_k\geqslant p_1$，即$p_k\geqslant 1-(1-p)^n$，有

$$1-p_1\geqslant(1-p)^n\Longrightarrow \ln(1-p_1)\geqslant n\ln(1-p),$$

所以 $$n\geqslant \ln(1-p_1)/\ln(1-p).$$

例14 在n次独立重复试验中，A在每次试验中发生的概率为0.3. 进行4次独立重复试验，若A一次不发生，则B也不发生；若A发生一次，则B发生的概率为0.6；若A发生两次或两次以上，则B一定发生. 求事件B发生的概率.

解 以A_0，A_1分别记A发生零次和一次的事件，以A_2记A发生两次和两次以上的事件，则A_0，A_1，A_2为一完备事件组，由二项概率公式

$$P(A_0)=C_4^0\times 0.3^0\times 0.7^4=0.2401,$$

$$P(A_1)=C_4^1\times 0.3\times 0.7^3=0.4116,$$

$$P(A_2)=1-P(A_0)-P(A_1)=0.3483.$$

又 $\quad P(B|A_0)=0,\quad P(B|A_1)=0.6,\quad P(B|A_2)=1,$

由全概率公式，得

$$P(B)=P(A_0)P(B|A_0)+P(A_1)P(B|A_1)+P(A_2)P(B|A_2)$$
$$=0+0.6\times0.4116+1\times0.3483=0.5953.$$

例 15 设某车间有 10 台同类型的设备，每台设备的电动机功率为 10 kW. 已知每台设备每小时实际开动 12 min，它们的使用是相互独立的. 因某种原因，这天供电部门只能给车间提供 50 kW 的电力. 问：这天这 10 台设备能正常运转的概率是多少？

解 依题意知，要求的是同时开动的设备不超过 5 台的概率.

本题可视为 10 重伯努利试验，$p=1/5$，故

$$P\{X\leqslant5\}=P\{X=0\}+P\{X=1\}+P\{X=2\}$$
$$+P\{X=3\}+P\{X=4\}+P\{X=5\}$$
$$=\sum_{i=0}^{5}C_n^i\left(\frac{1}{5}\right)^i\left(\frac{4}{5}\right)^{n-i}=0.994.$$

例 16（Banach 火柴盒问题） 某人带有两盒火柴，每盒有 n 根. 每次使用时，任取一盒中的一根，求：

(1) 发现一盒已空、另一盒恰剩 r 根的概率；

(2) 一盒取出最后一根时另一盒恰剩 r 根的概率.

解 以 B_1,B_2 记火柴取自不同两盒的事件，则有 $P(B_1)=P(B_2)=1/2$.

(1) 发现一盒已空、另一盒恰剩 r 根，说明已取了 $2n-r$ 次. 设 n 次取自 B_1 盒（已空），$n-r$ 次取自 B_2 盒，第 $2n-r+1$ 次拿起 B_1 盒，发现已空.

把取 $2n-r$ 次火柴视为 $2n-r$ 重伯努利试验，把取自 B_1 盒作为“成功”，则所求概率为

$$p_k=2\times C_{2n-r}^n\times(1/2)^n\times(1-1/2)^{2n-r-n}\times1/2=C_{2n-r}^n/2^{2n-r}.$$

式中 2 反映 B_1 与 B_2 盒的对称性（即也可以是 B_2 盒先取空），1/2 反映最后一次取 B_1（或 B_2）盒的概率.

（2）只需将总次数改为 $2n-r-1$，最后一次（即第 $2n-r$ 次取自 B_1（或 B_2）盒，将在 B_1 中共取 $n+1$ 次改为取 n 次即可. 故

$$p_k=C_{2n-r-1}^{n-1}/2^{2n-r-1}.$$

硕士研究生入学试题分析

一、本章考试要求（摘自研究生入学考试大纲，下同）

1. 了解样本空间（基本事件空间）的概念，理解随机事件的概念，掌握事件的关系与运算.

2. 理解概率、条件概率的概念，掌握概率的基本性质，会计算古典型概率和几何型概率，掌握概率的加法公式、减法公式、乘法公式、全概率公式以及贝叶斯公式.

3. 理解事件独立性的概念，掌握用事件独立性进行概率计算；理解独立重复试验的概念，掌握计算有关事件概率的方法.

二、本章重点内容

随机事件的关系与运算，古典型概率问题、几何型概率问题、条件概率问题，全概率公式与贝叶斯公式的应用、独立性与伯努利概型的应用.

（一）随机事件的概率

1. 将一枚硬币独立地掷两次，引进事件：$A_1=\{$掷第一次出现正面$\}$，$A_2=\{$掷第二次出现正面$\}$，$A_3=\{$正、反面各出现一次$\}$，$A_4=\{$正面出现两次$\}$，则事件（　　）.

（A）A_1,A_2,A_3 相互独立；　　（B）A_2,A_3,A_4 相互独立；

（C）A_1,A_2,A_3 两两独立；　　（D）A_2,A_3,A_4 两两独立.

（2003 年三）

解　选(C).（A）、（B）显然不成立，（D）也不成立. A_1,A_2 相互独立，A_1,A_3 相互独立，A_2,A_3 也相互独立，故(C)成立.

2. 对于任意二事件 A 和 B，（　　）.

(A) 若 $AB \neq \varnothing$,则 A,B 一定独立;

(B) 若 $AB \neq \varnothing$,则 A,B 有可能独立;

(C) 若 $AB \neq \varnothing$,则 B,B 一定独立;

(D) 若 $AB \neq \varnothing$,则 A,B 一定不独立. (2003 年四)

解 选(B).(C)显然不成立.

例如,当 $A=\Omega$(样本空间),且 $A \supset B$ 时,$P(AB)=P(B)=P(B)P(\Omega)=P(A)P(B)$,即 A,B 独立;但当 $A \neq \Omega$ 时,$P(AB) \neq P(A)P(B)$,即 A,B 不独立.故(A)、(D)也不成立.

3. 在电炉上安装了4个温控器,其显示温度的误差是随机的.在使用过程中,只要有2个温控器显示的温度不低于临界温度 t_0,电炉就断电.以 E 表示事件"电炉断电",设 $T_{(1)} \leqslant T_{(2)} \leqslant T_{(3)} \leqslant T_{(4)}$ 为4个温控器显示的按递增顺序排列的温度值,则事件 E 等于事件(　　).

(A) $\{T_{(1)} \geqslant t_0\}$;　　(B) $\{T_{(2)} \geqslant t_0\}$;

(C) $\{T_{(3)} \geqslant t_0\}$;　　(D) $\{T_{(4)} \geqslant t_0\}$. (2000 年三、四)

解 选(C),因为 $\{T_{(3)} \geqslant t_0\}$ 等价于有2个温控器显示的温度不低于临界温度.

4. 当事件 A,B 同时发生时,事件 C 必发生,则(　　).

(A) $P(C)=P(AB)$;

(B) $P(C)=P(A+B)$;

(C) $P(C) \leqslant P(A)+P(B)-1$;

(D) $P(C) \geqslant P(A)+P(B)-1$. (1992 年四、五)

解 选(D).因为

$$P(AB)=P(A)+P(B)-P(A+B),$$

即
$$P(AB) \geqslant P(A)+P(B)-1,$$

所以
$$P(C) \geqslant P(AB) \geqslant P(A)+P(B)-1.$$

5. 设 A 和 B 是两个概率不为零的不相容事件,则下列结论肯定正确的是(　　).

(A) $\overline{A}$ 与 $\overline{B}$ 不相容;　　(B) $\overline{A}$ 与 $\overline{B}$ 相容;

(C) $P(AB)=P(A)P(B)$； (D) $P(A-B)=P(A)$.

(1991 年四、五)

解 选(D).(A)和(C)显然不成立.当$A+B=\Omega$且A与B互斥时,$\overline{A}$与$\overline{B}$也互斥,故(B)也不成立.

6. 以A表示"甲种产品畅销,乙种产品滞销",则其对立事件$\overline{A}$为( ).

(A) 甲种产品滞销,乙种产品畅销；

(B) 甲、乙两种产品均畅销；

(C) 甲种产品畅销；

(D) 甲种产品滞销或乙种产品畅销. (1989 年四)

解 选(D).设$B=\{$甲种产品畅销$\}$,$C=\{$乙种产品畅销$\}$,则

$$\overline{A}=\overline{B\overline{C}}=\overline{B}+\overline{\overline{C}}=\overline{B}+C.$$

7. 对于任意两事件A和B,有$P(A-B)=$( ).

(A) $P(A)-P(B)$； (B) $P(A)-P(B)+P(AB)$；

(C) $P(A)-P(AB)$； (D) $P(A)+P(B)-P(AB)$.

(1987 年五)

解 选(C).作文氏图即可得知.

8. 已知A,B两个事件满足条件$P(AB)=P(\overline{A}\overline{B})$,且$P(A)=p$,则$P(B)=$________. (1994 年一)

解 因为

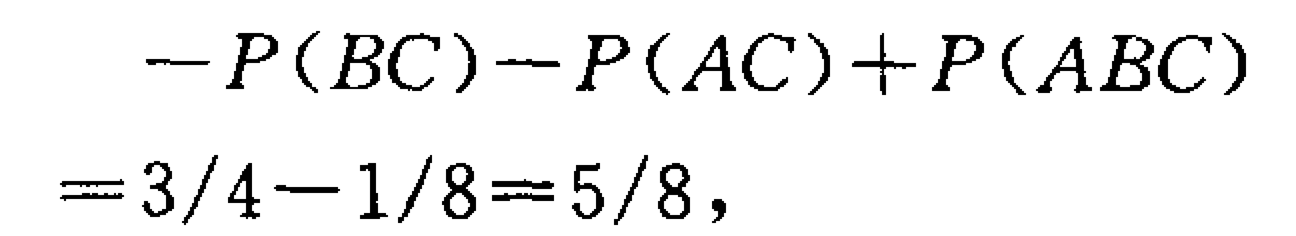
$$\begin{aligned}P(AB)&=P(\overline{A}\overline{B})=P(\overline{A\cup B})=1-P(A\cup B)\\&=1-[P(A)+P(B)-P(AB)]\\&=1-P(A)-P(B)+P(AB),\end{aligned}$$

所以 $$P(B)=1-P(A)=1-p.$$

9. 设$P(A)=P(B)=P(C)=1/4$,$P(AB)=P(BC)=0$,$P(AC)=1/8$,则A,B,C三事件中至少出现一个的概率为________. (1992 年五)

解 因为$P(AB)=P(BC)=0$,所以$P(ABC)=0$.于是

$$P(A+B+C)=P(A)+P(B)+P(C)-P(AB)$$



$$-P(BC)-P(AC)+P(ABC)$$
$$=3/4-1/8=5/8,$$

10. 已知 $P(A)=P(B)=P(C)=1/4, P(AB)=0, P(AC)=P(BC)=1/6$，则事件 A, B, C 全不发生的概率为________.

(1992 年一)

解 因为 $P(A+B+C)=1-P(A)-P(B)-P(C)+P(AB)+P(AC)+P(BC)=1-3/4+1/3=5/12$，所以

$$P(\overline{A}\overline{B}\overline{C})=1-5/12=7/12.$$

11. 将C,C,E,E,I,N,S 等7个字母随机地排成一排，那么恰好排成英文单词SCIENCE 的概率为________. (1992 年四、五)

解 $p=P_1^1P_2^1P_1^1P_2^1P_1^1P_1^1P_1^1/P_7^7=4/7!=1/1260.$

12. 随机地向半圆 $0<y<\sqrt{2ax-x^2}$（a 为正常数）内掷一点，点落在半圆内任何区域的概率与区域的面积成正比，则原点和该点的连线与 x 轴的夹角小于 $\pi/4$ 的概率为________. (1991 年一)

解 利用几何型概率求解. 图1.15 中半圆面积为 $\frac{1}{2}\pi a^2$，阴影部分面积为 $\frac{\pi}{4}a^2+\frac{1}{2}a^2$，故所求概率为

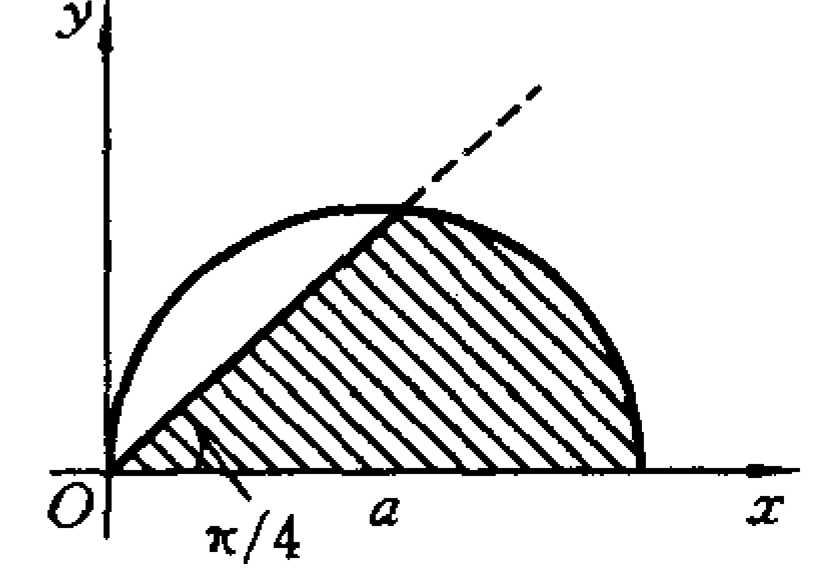


图1.15

$$\left(\frac{\pi}{4}a^2+\frac{1}{2}a^2\right)\Big/\left(\frac{\pi}{2}a^2\right)=\frac{1}{2}+\frac{1}{\pi}.$$

13. 设随机事件 A, B 及其和事件 $A\cup B$ 的概率分别为0.4, 0.3和0.6. 若 $\overline{B}$ 表示 B 的对立事件，那么积事件 $A\overline{B}$ 的概率 $P(A\overline{B})=$________.

(1990 年一)

解 因为 $P(A+B)=P(A)+P(B)-P(AB)$，所以

$$P(AB)=0.4+0.3-0.6=0.1,$$
$$P(A\overline{B})=P(A)-P(AB)=0.4-0.1=0.3.$$

14. 若在区间(0,1)内任取两个数，则事件{两数之和小于6/5}的概率为________. (1988 年一)

解 用几何型概率求解(见图1.16).所求概率为

$$p=1\times1-S_{\triangle BCD}=1-\frac{1}{2}\times\frac{4}{5}\times\frac{4}{5}=1-\frac{8}{25}=\frac{17}{25}.$$

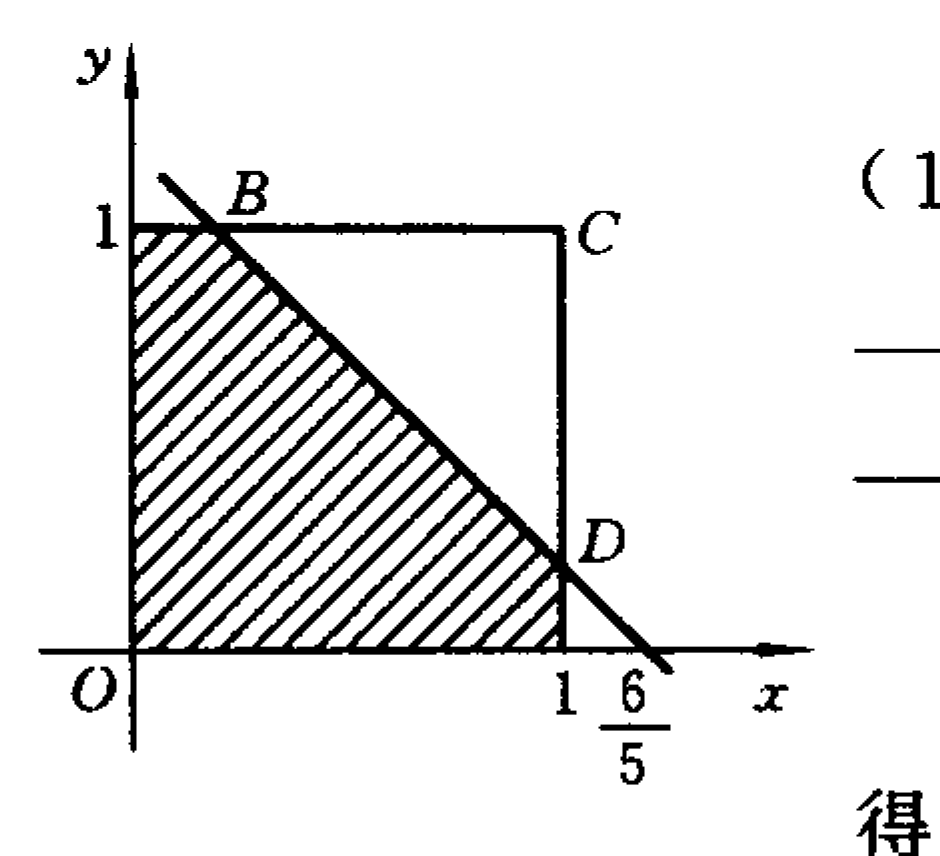


图1.16

15. 设 $P(A)=0.4, P(A+B)=0.7$. (1) 若 A,B 互不相容,则 $P(B)=$ ________;(2) 若 A,B 相容,则 $P(B)=$ ________. (1988年四)

解 (1) 若 $P(AB)=0$,则由

$$P(A+B)=P(A)+P(B),$$

得 $$P(B)=0.3.$$

(2) 若 $P(AB)\neq0$,则当 A,B 独立时,有 $P(A+B)=P(A)+P(B)-P(A)P(B)$,所以 $P(B)=0.5$.

16. 考虑一元二次方程 $x^2+Bx+C=0$,其中 B,C 分别为将一枚骰子接连掷两次先后出现的点数,求该方程有实根的概率 p 和有重根的概率 q. (1996年四)

解 组成 (B,C) 的事件有 $(1,1),\cdots,(6,6)$ 等36个,要使方程有实根,应有 $B^2-4C\geqslant0$;要使方程有重根,应有 $B^2=4C$.列表分析如表1.1所示,故

$$p=19/36,\quad q=1/18.$$

表1.1

B	1	2	3	4	5	6	$\sum$
$B^2\geqslant4C$	0	1	2	4	6	6	19
$B^2=4C$	0	1	0	1	0	0	2

17. 设 A,B 为两个事件,且 $P(A)=0.7, P(A-B)=0.3$,求 $P(\overline{AB})$. (1991年)

解 因为

$$P(AB)=P(A)-P(A-B)=0.7-0.3=0.4,$$

所以 $$P(\overline{AB})=1-P(AB)=0.6.$$

18. 从0,1,2,…,9等10个数字中任意选出3个不同的数字，试求下列事件的概率：

A_1={3个数字中不含0和5}；

A_2={3个数字中不含0或5}. （1990年四、五）

解 任选3个不同元素的组合有C_{10}^3种，不含0和5的组合有C_8^3种，不含0或5的组合有$2C_9^3-C_8^3$种，故

$$P(A_1)=C_8^3/C_{10}^3=7/15,$$

$$P(A_2)=(2C_9^3-C_8^3)/C_{10}^3=14/15.$$

19. 若事件A,B,C满足$A+C=B+C$，问：$A=B$是否成立？

（1988年四）

解 不一定. 如$A=\{1,2,3,4,5\}$，$B=\{1,2,3\}$，$C=\{4,5\}$，则$A+B=B+C$，但$A\neq B$.

20. 对于任意两事件A和B，与$A\cup B=B$不等价的是（　　）.

(A) $A\subset B$； (B) $\overline{B}\subset\overline{A}$；

(C) $A\overline{B}=\varnothing$； (D) $\overline{A}B=\varnothing$. （2001年四）

解 选(D). 因为$A\cup B=B$，即$A\subset B$，则$\overline{B}\subset\overline{A}$，$A\overline{B}=\varnothing$，所以不等价的是(D).

（二）条件概率与事件的独立性

1. 设A,B是两随机事件，且$0<P(A)<1$，$P(B)>0$，$P(B|A)=P(B|\overline{A})$，则必有（　　）.

(A) $P(A|B)=P(\overline{A}|B)$； (B) $P(A|B)\neq P(\overline{A}|B)$；

(C) $P(AB)=P(A)P(B)$； (D) $P(AB)\neq P(A)P(B)$.

（1998年一）

解 选(C). 因为$P(B|A)=P(B|\overline{A})$，所以

$$P(AB)/P(A)=P(\overline{A}B)/P(\overline{A}),$$

即
$$\frac{P(AB)}{P(A)}=\frac{P(\overline{A}B)}{P(\overline{A})}=\frac{P(B)-P(AB)}{1-P(A)},$$

$$P(AB)-P(AB)P(A)=P(A)P(B)-P(A)P(AB),$$

因此
$$P(AB)=P(A)P(B).$$

2. 已知 $0<P(B)<1$，且 $P[(A_1+A_2)|B]=P(A_1|B)+P(A_2|B)$，则下列选项成立的是(　　)：

(A) $P[(A_1+A_2)|\overline{B}]=P(A_1|\overline{B})+P(A_2|\overline{B})$；

(B) $P(A_1B+A_2B)=P(A_1B)+P(A_2B)$；

(C) $P(A_1+A_2)=P(A_1|B)+P(A_2|B)$；

(D) $P(B)=P(A_1)P(B|A_1)+P(A_2)P(B|A_2)$. (1996 年四)

解 选(B). 因为

$$P[(A_1+A_2)|B]=P[(A_1+A_2)B]/P(B),$$

又
$$\begin{aligned}P(A_1|B)+P(A_2|B)&=P(A_1B)/P(B)+P(A_2B)/P(B)\\&=[P(A_1B)+P(A_2B)]/P(B)\\&=P[(A_1+A_2)B]/P(B),\end{aligned}$$

所以 $P[(A_1+A_2|B]=P(A_1|B)+P(A_2|B)$ 与(B)等价.

3. 设 A,B 为任意两事件，且 $A\subset B$，$P(B)>0$，则下列选项必然成立的是(　　).

(A) $P(A)<P(A|B)$；　　(B) $P(A)\leqslant P(A|B)$；

(C) $P(A)>P(A|B)$；　　(D) $P(A)\geqslant P(A|B)$.

(1996 年五)

解 选(B). 当 $B\neq\Omega$ 时，$P(A)<P(A|B)$；当 $B=\Omega$ 时，$P(A)=P(A|B)$.

4. 设 A,B,C 三个事件两两独立，则 A,B,C 相互独立的充要条件是(　　).

(A) A 与 BC 独立；　　(B) AB 与 $A\cup C$ 独立；

(C) AB 与 AC 独立；　　(D) $A\cup B$ 与 $A\cup C$ 独立.

(2000 年四)

解 选(A). 若 A,B,C 相互独立，则

$$\begin{aligned}P(ABC)&=P(A)P(B)P(C)=P(A)[P(B)P(C)]\\&=P(A)P(BC),\end{aligned}$$

即 A 与 BC 独立. 又若 A 与 BC 独立，则

$$P(ABC)=P(A)P(BC)=P(A)P(B)P(C).$$

5. 设 $0<P(A)<1, 0<P(B)<1, P(A|B)+P(\overline{A}|\overline{B})=1$，则（　　）.

(A) 事件 A 和 B 互不相容；　　(B) 事件 A 和 B 相互对立；

(C) 事件 A 和 B 互不独立；　　(D) 事件 A 和 B 相互独立.

（1994 年五）

解　选(D). 因为若 A,B 相互独立，则

$$P(A|B)=P(A),\quad P(\overline{A}|\overline{B})=P(\overline{A}),$$

所以
$$P(A|B)+P(\overline{A}|\overline{B})=P(A)+P(\overline{A})=1.$$

6. 设两两独立的三事件 A,B,C 满足条件 $ABC=\varnothing, P(A)=P(B)=P(C)<1/2$，且已知 $P(A\cup B\cup C)=9/16$，则 $P(A)=$________.

（1999 年一）

解　$P(A\cup B\cup C)=P(A)+P(B)+P(C)-P(AB)-P(AC)-P(BC)+P(ABC)=3P(A)-3[P(A)]^2+0=9/16$. 解此关于 $P(A)$ 的方程，得 $P(A)=1/4$ (3/4 舍去).

7. 实习生用一台机器接连制造三个同种零件，第 i $(i=1,2,3)$ 个零件的不合格率 $p_i=1/(1+i)$. 以 X 表示三个零件中合格品的个数，则 $P\{X=2\}=$________.　（1996 年五）

解　因为 $\{X=2\}$ 表示三个零件中有两个合格的事件，由 {不合格，合格，合格}、{合格，不合格，合格}、{合格，合格，不合格} 三个事件组成，且各零件的合格与否显然相互独立，所以

$$P(X=2)=\frac{1}{2}\times\left(1-\frac{1}{3}\right)\left(1-\frac{1}{4}\right)+\left(1-\frac{1}{2}\right)\left(1-\frac{1}{4}\right)\times\frac{1}{3}+\left(1-\frac{1}{2}\right)\left(1-\frac{1}{3}\right)\times\frac{1}{4}=\frac{11}{24}.$$

8. 设 10 件产品中有 4 件不合格品，从中任取两件，已知所取两件产品中有一件是不合格品，则另一件也是不合格品的概率 p 是________.

（1993 年五）

解　设 A_i 为第 i 件产品为不合格品事件，则两件产品中有一件不合格的概率为

$$P(A_1A_2)+P(A_1\overline{A}_2)+P(\overline{A}_1A_2)$$
$$=4/10\times3/9+4/10\times6/9+6/10\times4/9=2/3,$$

故 $$p=(4/10\times3/9)/(2/3)=1/5.$$

9. 设两个相互独立的事件 A 和 B 都不发生的概率为 $1/9$，A 发生 B 不发生的概率与 B 发生 A 不发生的概率相等，则 $P(A)=$ ________. (2000 年一)

解 因为 $P(A\overline{B})=P(\overline{A}B)$，即

$$P(A)-P(AB)=P(B)-P(AB),$$

所以 $$P(A)=P(B).$$

而 $$P(\overline{A}\overline{B})=P(\overline{A})P(\overline{B})=[1-P(A)][1-P(B)]$$
$$=[1-P(A)]^2,$$

即 $[1-P(A)]^2=1/9$，得 $1-P(A)=1/3$，$P(A)=2/3$.

10. 设工厂 A 和工厂 B 的产品的次品率分别为 1%和 2%，现从由 A 和 B 的产品分别占 60%和 40%的产品中随机抽取一件，发现是次品，则该次品属工厂 A 生产的概率是________.

(1996 年一)

解 记 $A=\{$工厂A 的产品$\}$，$B=\{$工厂B 的产品$\}$，已知 $P(A)=0.6$，$P(B)=0.4$，A，B 构成完备事件组，$P(次|A)=0.01$，$P(次|B)=0.02$，则由全概率公式，有

$$P(次)=P(A)P(次|A)+P(B)P(次|B)$$
$$=0.6\times0.01+0.4\times0.02=0.014.$$

故所求概率由贝叶斯公式得

$$P(A|次)=P(A)P(次|A)/P(次)=3/7.$$

11. 袋中有50个乒乓球，其中20个是黄球，30个是白球，今有两人依次随机地从袋中各取一球，取后不放回，则第二个人取得黄球的概率为________. (1997 年一)

解 第二个人取得黄球概率为

$$P(黄|二)=P(黄|一)P(黄|黄)+P(白|一)P(黄|白)$$
$$=20/50\times19/49+30/50\times20/49=2/5.$$

12. 假设一批产品中一、二、三等品分别占60%,30%,10%,从中随意取出一件,结果不是三等品,则取到的是一等品的概率为________. (1994年五)

解 以 A_i $(i=1,2,3)$ 记一、二、三等品事件,则 $P(A_1)+P(A_2)=0.9$,所求概率为

$$p=P(A_1)/[P(A_1+A_2)]=0.6/0.9=2/3.$$

13. 设在三次独立试验中事件 A 出现的概率相等,若已知 A 至少出现一次的概率等于19/27,则 A 在一次试验中出现的概率为________. (1988年一)

解 因为 $P\{k\geqslant1\}=19/27,n=3$,所以由

$$P\{k\geqslant1\}=1-P\{k=0\}=1-(1-19/27)^3$$

可得 $$p=1-\sqrt[3]{1-19/27}=1-2/3=1/3.$$

14. 设有来自三个地区的、分别为10名、15名和25名考生的报名表,其中女生的报名表分别为3份、7份和5份.随机地取一个地区的报名表,从中先后抽出两份.

(1) 求先抽到的一份是女生表的概率;

(2) 已知后抽到的一份是男生表,求先抽到的一份是女生表的概率. (1998年三)

解 以 B_i $(i=1,2,3)$记第 i 个地区考生表的事件,以 A_j $(j=1,2)$记第 j 次取到男生表的事件,则 B_1,B_2,B_3 为一完备事件组,且

$$P(B_i)=1/3\ (i=1,2,3),\quad P(A_1|B_1)=7/10,$$
$$P(A_1|B_2)=8/15,\quad P(A_1|B_3)=20/25.$$

$$(1)\ P(\overline{A}_1)=\sum_{i=1}^{3}P(B_i)P(\overline{A}_1|B_i)$$
$$=1/3\times(3/10+7/15+5/25)=29/90.$$

(2) 因为 $P(A_2|B_1)=7/10,P(A_2|B_2)=8/15,P(A_2|B_3)=20/25,P(\overline{A}_1A_2|B_1)=7/30,P(\overline{A}_1A_2|B_2)=8/30,P(\overline{A}_1A_2|B_3)=5/30$,所以

$$P(A_2)=\sum_{i=1}^{3}P(B_i)P(A_2|B_i)=\frac{1}{3}\left(\frac{7}{10}+\frac{8}{15}+\frac{20}{25}\right)=\frac{61}{90}.$$

$$P(\overline{A}A_2)=\sum_{i=1}^{3}P(B_i)P(\overline{A}_1\overline{A}_2|B_i)=\frac{1}{3}\left(\frac{7}{30}+\frac{8}{30}+\frac{5}{30}\right)=\frac{2}{9}.$$

$$P(\overline{A}_1|A_2)=P(\overline{A}_1A_2)/P(A_2)=\frac{2}{9}\Big/\frac{61}{90}=\frac{20}{61}.$$

15. 一射手对同一目标独立地进行4次射击，若至少击中一次的概率为80/81，则该射手的命中率是多少？ （1990年四）

解 因为一次也没击中的概率为

$$(1-p)^4=1-80/81=1/81,$$

所以 $$p=1-\sqrt[4]{1/81}=1-1/3=2/3.$$

16. 玻璃杯成箱出售，每箱20只．假设各箱含0，1，2只残次品的概率分别为0.8，0.1，0.1．一顾客欲购一箱玻璃杯，在购买时，售货员随意取一箱，顾客开箱随机地察看4只，若无残次品，则买下该箱玻璃杯；否则退回．试求：

（1）顾客买下该箱的概率 α；

（2）在顾客买下的一箱中，确实没有残次品的概率 β.

（1988年四、五）

解 以 A_i $(i=0,1,2)$ 记一箱中有 i 只残次品事件，以 B 记顾客买下该箱的事件，则

$$P(A_0)=0.8,\quad P(A_1)=P(A_2)=0.1,\quad P(B|A_0)=1,$$

$$P(B|A_1)=C_{19}^4/C_{20}^4=4/5,\quad P(B|A_2)=C_{18}^4/C_{20}^4=12/19.$$

于是 $$\alpha=P(B)=\sum_{i=0}^{2}P(A_i)P(B|A_i)$$

$$=0.8\times1+4/5\times0.1+12/19\times0.1\approx0.943.$$

$$\beta=P(B|A_0)=P(A_0B)/P(B)=P(A_0)P(B|A)/P(B)$$

$$=0.8\times1/0.943=0.848.$$

17. 有两个箱子，第一个箱子里有3个白球，2个红球；第二个箱子里有4个白球，4只红球．现从第一个箱子里随机地取一个球

放到第二个箱子里，再从第二个箱子里取出一个球，此球是白球的概率是________.已知从上述第二个箱子里取出的球是白球，则从第一个箱子里取出的球是白球的概率是________. (1987 年一)

解 以 A_i $(i=1,2)$记从第 i 个箱子里取到白球的事件，则

$$P(A_1)=3/5,\quad P(A_2|A_1)=5/9,\quad P(A_2|\overline{A}_1)=4/9.$$

由全概率公式，有

$$\begin{aligned}P(A_2)&=P(A_1)P(A_2|A_1)+P(\overline{A}_1)P(A_2|\overline{A}_1)\\&=3/5\times5/9+2/5\times4/9=23/45,\end{aligned}$$

故
$$\begin{aligned}P(A_1|A_2)&=P(A_1A_2)/P(A_2)=(3/5\times5/9)/(23/45)\\&=15/23.\end{aligned}$$

18. 一批产品共有10个正品和2个次品，任意抽取两次，每次抽一个，抽出后不再放回，则第二次抽出的是次品的概率为________. (1993 年一)

解 以 A_i $(i=1,2)$记第 i 次取得次品的事件，则

$$\begin{aligned}P(A_2)&=P(A_1)P(A_2|A_1)+P(\overline{A}_1)P(A_2|\overline{A}_1)\\&=2/12\times1/11+10/12\times2/11=22/132=1/6.\end{aligned}$$

19. 电源电压在不超过200 V、200～240 V 和超过240 V 三种情况下，元件损坏的概率分别为0.1，0.001和0.2.设电源电压服从正态分布，$X\sim N(220,25^2)$，求：

(1) 元件损坏的概率 α；

(2) 元件损坏时，电压在200～240 V 间的概率 β.

解
$$\begin{aligned}P\{X\leqslant200\}&=P\left\{\frac{X-220}{25}\leqslant\frac{200-220}{25}\right\}=\Phi(-0.8)\\&=1-\Phi(0.8)=0.2118,\end{aligned}$$

$$\begin{aligned}P\{200<X\leqslant240\}&=P\left\{\frac{200-220}{25}<\frac{X-220}{25}\leqslant\frac{240-220}{25}\right\}\\&=2\Phi(0.8)-1=0.5763,\end{aligned}$$

$$P\{X>240\}=1-P\{X\leqslant240\}=1-\Phi(0.8)=0.2119.$$

由全概率公式知，元件损坏的概率为

$$\alpha=0.1\times0.2118+0.001\times0.5763+0.2\times0.2119$$

$=0.0641.$

由贝叶斯公式知，元件损坏时，电压在200～240 V的概率为

$$\beta=0.5763\times0.001/0.0641=0.0090.$$

20. 设A,B是任意两事件，其中A的概率不等于0和1，证明$P(B|A)=P(B|\overline{A})$是事件A和B独立的充分必要条件.

（2002年四）

证 由于A的概率不等于0和1，故题中两个条件概率都存在.

(1) 必要性 由事件A与B独立，知事件$\overline{A}$与B也独立，因此

$$P(B|A)=P(B),\quad P(B|\overline{A})=P(B),$$

从而

$$P(B|A)=P(B|\overline{A}).$$

(2) 充分性 由$P(B|A)=P(B|\overline{A})$可见

$$\frac{P(AB)}{P(A)}=\frac{P(\overline{A}B)}{P(\overline{A})}=\frac{P(B)-P(AB)}{1-P(A)},$$

故

$$P(AB)[1-P(A)]=P(A)P(B)-P(A)P(AB),$$

即

$$P(AB)=P(A)P(B).$$

第二章　随机变量及其概率分布

第一节　随机变量及其分布函数

主 要 内 容

1. 随机变量

设Ω是随机试验E的样本空间，若对每一个样本点$\omega\in\Omega$，有一个实数$X(\omega)$与之对应，这样得到的一个定义在样本空间上的实值单值函数$X=X(\omega)$，称为一个随机变量X.

2. 分布函数

设 X 为一个随机变量，x 是任意实数，则称函数 $F(x)=P\{X\leqslant x\}$为X 的分布函数，也称之为累积概率函数.

3. 分布函数的性质

(1) $F(x)$是一个不减函数，即若$x_1\leqslant x_2$，则$F(x_1)\leqslant F(x_2)$；

(2) $0\leqslant F(x)\leqslant 1$，且$\lim\limits_{x\to-\infty}F(x)=0$，$\lim\limits_{x\to+\infty}F(x)=1$；

(3) $F(x)$右连续，即$F(x+0)=F(x)$.

疑 难 解 析

1. 随机变量与普通函数有何区别？引入随机变量有何意义？

答　随机变量是一个单值实值函数. 它是对随机试验E的样本空间Ω的每一样本点$\omega\in\Omega$，定义一个实数而得到的一个函数. 它与普通函数的区别是：(1)定义域是样本空间，不是实轴上的区

间；(2)随机变量 X 的值在试验前是不确定的，按统计规律性给出取值的概率，因而具有随机性，而普通函数的取值是由对应法则 f 确定的.

引入随机变量是为了研究随机现象的统计规律性. 我们将形形色色的样本空间和样本点统一化、数量化，使之与实轴上的一个集合或者点对应起来，就可以用微积分的理论与方法对随机试验与随机事件的概率进行数学推理与计算，从而完成对随机试验结果的规律性的研究. 因而，随机变量的引入具有重要的意义.

2. 随机变量的分布函数有什么意义？

答 分布函数 $F(x)=P\{X\leqslant x\}$ 反映了随机变量 X 的取值不大于实数 x 的概率，故又称累积概率函数. X 在实轴任意区间 $(x_1,x_2]$ 上的概率也可以用 $F(x)$ 来表示，即 $P\{x_1<X\leqslant x_2\}=F(x_2)-F(x_1)$. 因此，掌握了随机变量 X 的分布函数，就了解了随机变量 X 在 $(-\infty,+\infty)$ 上的概率分布. 可以说，分布函数完整地描述了随机变量的统计规律性.

分布函数与随机变量不同，它是一个普通函数，有定义域 $(-\infty,+\infty)$，有对应法则 $F(x)=P\{X\leqslant x\}$，因此，非常便于用微积分的理论与方法进行研究、分析与计算. 概率论与数理统计就是借助随机变量与分布函数来全面研究与认识随机现象的统计规律性的.

3. 分布函数的左、右连续定义有什么区别？

答 分布函数有两种定义方法. 目前，俄罗斯等东欧国家使用的是左连续定义，其它许多国家使用的是右连续定义. 两者的区别在于：左连续定义定义 $F(x)=P\{X<x\}$，而右连续定义定义 $F(x)=P\{X\leqslant x\}$. 这就决定了在计算 $F(x)$ 或 $P\{x_1<X<x_2\}$ 时，端点 $X=x_1$ 或 $X=x_2$ 的概率是否计算在内. 当 X 为离散型随机变量时，$P\{X=x_1\}$ 或 $P\{X=x_2\}$ 可能不为零，因此左连续和右连续的分布函数或概率可能不同. 对此，读者要予以充分注意，以免出错. 当 X 为连续型随机变量时，因为一点上的概率为零，所以左连续定义与右连续定义是一样的.

4. 不同的随机变量,它们的分布函数是否一定不同?

答 否,可能相同. 例如,掷一枚均匀的硬币,可以令

$$X_1=\begin{cases}1, & 正面朝上,\\ -1, & 反面朝上,\end{cases} \qquad X_2=\begin{cases}1, & 反面朝上,\\ -1, & 正面朝上.\end{cases}$$

显然,X_1 与 X_2 是两个不同的随机变量,因为它们有不同的对应法则,但它们的分布函数相同,即

$$F(x)=\begin{cases}0, & x<-1,\\ 1/2, & -1\leqslant x<1,\\ 1, & x\geqslant 1.\end{cases}$$

方法、技巧与典型例题分析

在验证某一函数是否可以作为分布函数时,一定要从定义 $F(x)=P\{X\leqslant x\}$ 出发,考察所讨论函数是否具备分布函数的性质,确定是否为分布函数. 同时,求随机变量 X 的分布函数也要从定义 $F(x)=P\{X\leqslant x\}$ 出发,自左至右考察 $(-\infty,+\infty)$ 上的每一点 x. 当 $F(x)$ 是分段函数时,必须清楚表明 $F(x)$ 的不同定义域段和不同解析式,牢记分布函数是累积概率这一特性.

例 1 分析下列函数中哪个是随机变量 X 的分布函数.

(1) $F_1(x)=\begin{cases}0, & x<-2,\\ 1/2, & -2\leqslant x<0,\\ 2, & x\geqslant 0;\end{cases}$

(2) $F_2(x)=\begin{cases}0, & x<0,\\ \sin x, & 0\leqslant x<\pi,\\ 1, & x\geqslant \pi;\end{cases}$

(3) $F_3(x)=\begin{cases}0, & x<0,\\ x+1/2, & 0\leqslant x<1/2,\\ 1, & x\geqslant 1/2.\end{cases}$

解 (1) 不是分布函数,因为 $\lim\limits_{x\to+\infty}F_1(x)=2$,不符合 $0\leqslant F(x)\leqslant 1$.

(2) 不是分布函数，因为 $F_2(x)=\sin x$ 在 $(\pi/2,\pi)$ 是单调减少的，不是不减函数.

(3) 是分布函数，符合分布函数三条性质. 但 $F(x)$ 在 $x=0$ 与 $x=1/2$ 处不可导，且由此得出 $\int_{-\infty}^{x} F'(x)\mathrm{d}x \neq F(x)$. 由下节知，不存在概率密度函数，同时 $F(x)$ 图形也不是阶跃曲线，所以 $F(x)$ 既非连续型也非离散型随机变量的分布函数.

例 2 设 X 的分布函数为

$$F(x)=\begin{cases} b, & x<-1, \\ a, & -1\leqslant x<1, \\ 2/3-a, & 1\leqslant x<2, \\ a+b, & x\geqslant 2, \end{cases}$$

且 $P\{X=2\}=1/2$，求 a,b 和 X 的概率分布.

解 由题设知 $a+b=1$，$2/3-a=1/2$，故 $a=1/6$，$b=5/6$. 于是

$$P\{X=-1\}=1/6,\quad P\{X=1\}=1/3,\quad P\{X=2\}=1/2.$$

例 3 向直线上掷随机点，已知随机点落入区间 $H_1=(-\infty,0]$，$H_2=(0,1]$，$H_3=(1,+\infty]$ 的概率分别为 0.2，0.5，0.3，且随机点在 $(0,1]$ 上是均匀的. 设随机点落入区间 H_1,H_2,H_3 分别得 $0,x,1$ 分，以 X 记得分，求 X 的分布函数.

解 以 H_i 记随机点落入区间 H_i 的事件，则

$$P(H_1)=0.2,\quad P(H_2)=0.5,\quad P(H_3)=0.3,$$

于是

$$P\{X\leqslant x\mid H_1\}=\begin{cases} 0, & x<0, \\ 1, & x\geqslant 0, \end{cases}$$

$$P\{X\leqslant x\mid H_2\}=\begin{cases} 0, & x<0, \\ x, & 0\leqslant x<1, \\ 1, & x\geqslant 1, \end{cases}$$

$$P\{X<x\mid H_3\}=\begin{cases} 0, & x<1, \\ 1, & x\geqslant 1. \end{cases}$$

由全概率公式

$$F(x)=P\{X\leqslant x\}=\sum_{i=1}^{3}P(H_i)P\{X\leqslant x|H_i\},$$

得
$$F(x)=\begin{cases}0, & -\infty<x<0,\\ 0.2+0.5x, & 0\leqslant x<1,\\ 1, & 1\leqslant x<+\infty.\end{cases}$$

例 4 设随机变量 X 的分布函数如下：

$$F(x)=\begin{cases}1/(1+x^2), & x<\underline{\quad ①\quad},\\ \underline{\quad ②\quad}, & x\geqslant\underline{\quad ③\quad}.\end{cases}$$

试填上①,②,③项.

解 分布函数一定有 $\lim\limits_{x\to+\infty}F(x)=1$，故②项填1；又，$F(x)$是连续函数，显然有$\lim\limits_{x\to 0}\dfrac{1}{1+x^2}=1$，故①项填0，从而③项亦填0. 即

$$F(x)=\begin{cases}1/(1+x^2), & x<0,\\ 1, & x\geqslant 0.\end{cases}$$

例 5 设随机变量 X 的分布函数为

$$F(x)=A+B\arctan x\quad(-\infty<x<+\infty),$$

求：(1) A 与 B；(2) $P\{|X|<1\}$.

解 (1) 由 $\lim\limits_{x\to-\infty}F(x)=0$，$\lim\limits_{x\to+\infty}F(x)=1$，得

$$\lim_{x\to-\infty}(A+B\arctan x)=A-\pi B/2=0,$$
$$\lim_{x\to+\infty}(A+B\arctan x)=A+\pi B/2=1.$$

解得
$$A=1/2,\quad B=1/\pi.$$

(2)
$$P\{|x|<1\}=P\{-1<x<1\}=F(1)-F(-1)$$
$$=\left(\frac{1}{2}+\frac{1}{\pi}\times\frac{\pi}{4}\right)-\left(\frac{1}{2}-\frac{1}{\pi}\times\frac{\pi}{4}\right)=\frac{1}{2}.$$

例 6 掷一枚均匀的骰子，以 $X=i$ 记出现的点数为偶数的次数，求 X 的分布函数.

解 X 的取值 i 只有0,1两个值. 以 ω_j 记骰子出现 j $(j=1,2,\cdots,6)$点的事件，$P(\omega_j)=1/6$，所以

$$P\{X=0\}=P\{\omega_1\cup\omega_3\cup\omega_5\}=1/2,$$

$$P\{X=1\}=P\{\omega_2\cup\omega_4\cup\omega_6\}=1/2,$$

故
$$F(x)=\begin{cases}0, & x<0,\\ 1/2, & 0\leqslant x<1,\\ 1, & x\geqslant 1.\end{cases}$$

例 7 一均匀陀螺，在其圆周的半圈上均匀地刻上区间[0,1]上的数字，另半圈表示数字1. 旋转这陀螺，求陀螺停下时其圆周上触及桌面的点的刻度 X 的分布函数.

解 以 H_1 表示点落在表示 1 的半圈的事件，H_2 表示点在另半圈的事件，则

$$P(H_1)=P(H_2)=1/2,$$

$$P\{X\leqslant x\mid H_1\}=\begin{cases}1, & x\geqslant 1,\\ 0, & \text{其它},\end{cases}$$

$$P\{X\leqslant x\mid H_2\}=\begin{cases}0, & x<0,\\ x, & 0\leqslant x<1,\\ 1, & x\geqslant 1.\end{cases}$$

由全概率公式，知

$$F(x)=P\{X\leqslant x\}=\sum_{i=1}^{2}P(H_i)P\{X\leqslant x\mid H_i\}=\begin{cases}0, & x<0,\\ x/2, & 0\leqslant x<1,\\ 1, & x\geqslant 1.\end{cases}$$

分布函数既不是连续型的，也不是离散型的(见下节内容).

第二节　离散型随机变量及其概率分布

主 要 内 容

1. 离散型随机变量

(1) 如果随机变量 X 的取值是有限个或可列无限多个，则称

X 为离散型随机变量.

(2) 若 X 的所有可能取值 x_k 的概率

$$P\{X=x_k\}=p_k,\quad k=1,2,\cdots,$$

且满足 $p_k\geqslant 0,\sum\limits_{k=1}^{\infty}p_k=1$,则称上式为离散型随机变量 X 的分布律,或称之为概率分布、概率函数. 当用表格形式给出时,也可称之为分布列.

(3) 对离散型随机变量,若 $P\{X=x_k\}=p_k$,则

$$F(x)=\sum_{x_k\leqslant x}p_k.$$

此时,$F(x)$ 为一阶跃曲线,在每一 x_k 有一跃度.

2. 一些常用的离散型随机变量及其分布

(1) 0-1 分布(二点分布)　随机变量 X 只取 0 和 1 两个值,其分布律是

$$P\{X=k\}=p^k(1-p)^{1-k},\ k=0,1\quad(0<p<1).$$

只有两个结果的随机试验可以定义为 0-1 分布. 也可以把只有两类结果的随机试验定义为 0-1 分布.

(2) 二项分布　若随机变量 X 的可取值为 $k=0,1,\cdots,n$,其分布律是

$$P\{X=k\}=C_n^kp^k(1-p)^{n-k}\quad(0<p<1),$$

记为 $X\sim B(n,p)$,称 X 服从参数为 n,p 的二项分布.

二项分布是 n 重伯努利试验中事件 A 恰好发生 k 次的概率的分布.

(3) 泊松分布　若随机变量 X 的可取值为 $k=1,2,\cdots,n,\cdots$,其分布律是

$$P\{X=k\}=\frac{\lambda^k}{k!}\mathrm{e}^{-\lambda}\quad(\lambda>0,\text{常数}),$$

记为 $X\sim\pi(\lambda)$(或 $p(\lambda)$),称 X 服从参数为 λ 的泊松分布.

泊松分布又称为空间散布点子的几何模型. 当事件"流"满足平稳性、无后效性、普通性时,事件的概率服从泊松分布.

(4) 几何分布　若随机变量 X 的可取值为 $k=1,2,\cdots,n,\cdots$，其分布律是

$$P\{X=k\}=p(1-p)^{k-1}\quad(0<p<1),$$

记为 $X\sim G(p)$，称 X 服从参数为 p 的几何分布.

(5) 超几何分布　若随机变量 X 的可取值为 $k=0,1,\cdots,l$ $(l=\min\{M,N\})$，其分布律是

$$P\{X=k\}=\frac{C_M^k C_{N-M}^{n-k}}{C_N^n},$$

记为 $X\sim H(n,M,N)$，称 X 服从参数为 n,M,N 的超几何分布.

疑难解析

1. 试描述二项分布的性态.

答　二项分布 $X\sim B(n,p)$ 的概率 $P\{X=k\}$ 随着 k 的增大而增大，在达到最大值后，又随着 k 的增大而变小. 这是因为

$$\frac{P\{X=k+1\}}{P\{X=k\}}=\frac{(n-k)p}{(k+1)(1-p)}.$$

故由 $(n-k)p-(k+1)(1-p)=(n+1)p-1-k$ 知：当 $(n+1)p$ 为整数且当 $k=(n+1)p-1$ 或 $k=(n+1)p$ 时，$P\{X=k\}$ 取得最大值；当 $(n+1)p$ 不是整数且当 $k=[(n+1)p]$（取整）时，$P\{X=k\}$ 取得最大值.

当 $n<p/(1-p)$ 时，$P\{X=k\}$ 单调增加，$P\{X=n\}$ 为最大值；当 $n<(1-p)/p$ 时，$P\{X=k\}$ 单调减少，$P\{X=0\}$ 为最大值.

当 $p=0.5$ 时，二项分布是对称的；当 $p\neq0.5$ 时，二项分布是不对称的，且 n 越大，不对称性越不明显.

2. 试描述泊松分布的性态.

答　泊松分布 $X\sim\pi(\lambda)$ 的概率 $P\{X=k\}$ 先随着 k 的增大而增大，达到最大值后，又随着 k 的增大而减小. 这是因为

$$\frac{P\{X=k\}}{P\{X=k-1\}}=\frac{\lambda^k e^{-\lambda}(k-1)!}{\lambda^{k-1}e^{-\lambda}k!}=\frac{\lambda}{k}.$$

当$k<\lambda$时，有$P\{X=k\}>P\{X=k-1\}$. 故当λ为整数时，$k=\lambda$或$\lambda-1$，$P\{X=k\}$为最大值；当λ不是整数时，$k=[\lambda]$(取整)，$P\{X=k\}$为最大值.

当$\lambda<1$时，$P\{X=k\}$单调减少，$P\{X=0\}$为最大值.

泊松分布是不对称的，且λ越大，不对称性越不明显.

3. 超几何分布、二项分布和泊松分布之间有什么联系与区别?

答 先引入一个例子. 若一箱中有N件产品，其中有M件次品，一次抽取一件，共抽取n件，要求出取到的次品个数的分布.

若抽取是有放回的，则每次抽取时样本空间不变，抽取是独立的，次品数X服从参数为n,p的二项分布. 若抽取是无放回的，则继续抽取时样本空间改变，次品数X服从参数为n,M,N的超几何分布.

泊松分布是作为二项分布的近似而引入的. 假定事件"流"满足：(1)平稳性，即"流"的发生次数只与时间间隔Δt的长短有关，而与初始时刻无关；(2)无后效性，即任一时刻t_0前"流"的发生与t_0后"流"的发生无关；(3)普通性，即当时间间隔Δt很小时，"流"至多只发生一次. 这时事件的发生次数的概率分布服从泊松分布.

在计算概率的过程中，泊松分布有表可查，而超几何分布与二项分布的计算都较麻烦. 当n很大而p很小时，有以下近似公式：

$$\frac{C_M^k C_{N-M}^{n-k}}{C_N^M} \approx C_n^k p^k (1-p)^{n-k} \approx \frac{\lambda^k e^{-\lambda}}{k!} \quad (\lambda=np).$$

其中，第一个"≈"号，要求n很大，n/N较小，且取$p=n/N$；第二个"≈"号，n至少应不小于20，当$n\geqslant 50$时效果较好.

方法、技巧与典型例题分析

对离散型随机变量而言，通常有两类问题.

(1) 根据具体问题，求出随机变量X的分布律. 这时，首先要

明确X的取值范围，即X可以取哪些值；然后逐个计算$P\{X=x_i\}$，往往要借助古典型概率、条件概率、概率的运算性质和公式进行.

（2）利用离散型随机变量的分布律或分布函数求事件的概率.这时，要正确分析事件的组成与运算关系，求得准确的结果.

例1 同时掷两枚骰子，直到一枚骰子出现6点为止，求抛掷次数X的分布律.

解 先求每次掷出现6点的事件，以A,B分别记第一、二枚骰子出现6点的事件，则$P(A)=P(B)=1/6$，且A,B相互独立.以C记每次抛掷出现6点事件，则

$$\begin{aligned}P(C)&=P(A+B)=P(A)+P(B)-P(A)P(B)\\&=1/6+1/6-1/6\times1/6=11/36.\end{aligned}$$

故在k次试验中，第k次才出现6点的概率为

$$P\{X=k\}=(11/36)(1-11/36)^{k-1},\quad k=1,2,\cdots.$$

这是几何分布，$X\sim G(11/36)$.

例2 一盒中装有编号为1,2,…,5的5个球，现从中任取3个球，求被抽取的3个球的中间号码数X的分布律.

解 X可取值为2,3,4.当$X=k$时，另2个球中的1个在小于k的$k-1$个球中取，剩下1个球在大于k的$5-k$个球中取.故

$$P\{X=k\}=(C_{k-1}^1C_{5-k}^1)/C_5^3,\quad k=2,3,4,$$

即

X	2	3	4
p_k	0.3	0.4	0.3

例3 在n重伯努利试验中，每次试验中A发生的概率为p，以X记A发生k次所需进行的试验次数，求X的分布律.

解 X可取值为$k,k+1,\cdots,n$. $X=x$表示第x次试验一定成功，而在前$x-1$次试验中成功了$k-1$次（是二项分布），故分布律（称巴斯卡分布）为

$$P\{X=k\}=pC_{x-1}^{k-1}p^{k-1}(1-p)^{x-k}=C_{x-1}^{k-1}p^k(1-p)^{x-k}.$$

例4 设随机变量X的所有可能取值为整数1,2,…,10,又已

知 $P\{X=k\}$正比于k 的值，求 X 的分布律.

解 依题意，$P\{X=k\}=Ck$ (C 为常数)，于是

$$\sum_{k=1}^{10} Ck=55C=1 \Longrightarrow C=\frac{1}{55}.$$

故 X 的分布律为 $P\{X=k\}=k/55,\ k=1,2,\cdots,10.$

例 5 一箱中装有80 只管纱，其中棉纱48 只，腈纶纱32 只，现从箱中抽取两次，每次 1 只. 以 X 记可能抽到的棉纱的只数. 试写出在放回与不放回抽样下 X 的概率分布.

解 在放回抽样下，$X\sim B(2,48/80)$，故

$$P\{X=0\}=(32/80)^2=0.16,$$

$$P\{X=1\}=C_2^1\times(32/80)\times(48/80)=0.48,$$

$$P\{X=1\}=C_2^2\times(48/80)^2=0.36.$$

在不放回抽样下，$X\sim H(2,48,80)$，故

$$P\{X=0\}=C_{32}^2/C_{80}^2=0.157,$$

$$P\{X=1\}=C_{48}^1C_{32}^1/C_{80}^2=0.486,$$

$$P\{X=2\}=C_{48}^2/C_{80}^2=0.357.$$

可写出分布律如下：

X	0	1	2
放回抽样 p_k	0.16	0.48	0.36
不放回抽样 p_k	0.157	0.486	0.357

例 6 某运动员参加射箭比赛，共有 4 支箭，设其每支箭的命中率为 p，且各次射箭是相互独立的. 以 X 记直至命中为止所需射箭的次数，求 X 的概率分布.

解 X 可取值为 1,2,3,4. 记 $q=1-p$，则

$$P\{X=1\}=p,\quad P\{X=2\}=qp,\quad P\{X=3\}=q^2p,$$

$$P\{X=4\}=q^3(p+1-p)=q^3$$

(因为最后一箭可能射中也可能射不中，所以不等于 pq^3).

例 7 将 1～9 等 9 个数放入 3×3 的格子中，每格随机放入一

数. 设各列的最小值为 k_1,k_2,k_3，求 $X=\min\{k_1,k_2,k_3\}$ 的分布律.

解 事件的组成较为复杂，需要详细进行讨论.

依题意知，X 可取值为 3,4,…,7,9 个数放入 9 个格子的基本事件总数为 9!.

当 $T=3$ 时，k_1,k_2,k_3 分别为 1,2,3，有 $9\times6\times3$ 种放法，4～9 放入其余 6 格有 6!种放法，故

$$P(X=3)=(9\times6\times3\times3!)/9!=9/28.$$

当 $T=4$ 时，k_1,k_2,k_3 中含 1 和 4，另一数为 2 或 3. 当为 2 时，3 只能与1 或4 同列，有4 种放法，其余5 个数有5 个位置共5!种放法. 当为3 时，2 只能与1 同列，有2 种放法，其余5 个数有5!种放法，故

$$P\{X=4\}=(9\times6\times3)(4\times5!+2\times5!)/9!=9/28.$$

当 $T=5$ 时，k_1,k_2,k_3 中含1 和5，另一数为2 或3 或4. 同理可算得

$$P\{X=5\}=(9\times6\times3)(4\times3\times4!+4\times4!+2\times4!)/9!=6/28.$$

当 $X=6$ 时，k_1,k_2,k_3 中含1 与6，另一个数为2 或3 或4. 同理可算得

$$P\{X=6\}=(9\times6\times3)(4\times3\times3!+3\times2\times2\times3!+2\times2\times3!)/9!=3/28.$$

当 $X=7$ 时，k_1,k_2,k_3 中含1 与7，另一个数为2 或3 或4. 同理可算得

$$P\{X=7\}=(9\times6\times3)(4!\times2!+3\times2\times2\times2!+2\times2\times2!)/9!=1/28.$$

因此分布律为

X	3	4	5	6	7
p_k	9/28	9/28	6/28	3/28	1/28

例 8 求长度为 l 的棉纱上所含疵点数的概率分布.

解 X 可取值为正整数 0,1,2,….

将棉纱一端取作原点，把区间等分为长为 l/n 的 n 个半开区间. 设在每一小区间上出现一个以上疵点的概率为零，出现一个疵点的概率 $p_n=\mu\Delta l$. 且各小段上出现疵点相互独立.

以 A 记一段上出现一个疵点事件，则

$$P(A)=p_n=\mu\Delta l=\mu l/n, \quad P(\overline{A})=1-p_n.$$

观察 n 个区间上出现疵点是 n 重伯努利试验，因此，有

$$P\{X=k\}=C_n^k p_n^k(1-p_n)^{n-k},$$

将$[0,l]$细分，令 $n\to\infty$，有

$$P\{X=k\}=\lim_{n\to\infty}C_n^k p_n^k(1-p_n)^{n-k}.$$

记 $\lambda=\mu l$，则 $p_n=\lambda/n$，有

$$\begin{aligned}&C_n^k p_n^k(1-p_n)^{n-k}\\&=\frac{n!}{k!\ (n-k)!}\left(\frac{\lambda}{n}\right)^k\left(1-\frac{\lambda}{n}\right)^{n-k}\\&=\frac{\lambda^k}{k!}\cdot\frac{n(n-1)\cdots(n-k+1)}{n^k}\cdot\left(1-\frac{\lambda}{n}\right)^n\Big/\left(1-\frac{\lambda}{n}\right)^k.\end{aligned}$$

而

$$\lim_{n\to\infty}\frac{n(n-1)\cdots(n-k+1)}{n^k}=1,$$

$$\lim_{n\to\infty}\left(1-\frac{\lambda}{n}\right)^n=e^{-\lambda}, \quad \lim_{n\to\infty}\left(1-\frac{\lambda}{n}\right)^k=1,$$

得

$$P\{X=k\}=\lim_{n\to\infty}C_n^k p_n^k(1-p_n)^k=\frac{\lambda^k}{k!}e^{-\lambda}.$$

所以，棉纱上的疵点数 X 服从泊松分布.

例9 设随机变量 X 的分布律为 $P\{X=k\}=k/15$，$k=1,2,\cdots,5$，则在概率 $P\{1/2\leqslant X\leqslant 5/2\}$，$P\{1\leqslant X\leqslant 2\}$，$P\{0<X<3\}$，$P\{X=1$ 和 $X=2\}$，$P\{X=3\}$ 中，值等于 1/5 的有(　　)个.

(A) 2;　　(B) 3;　　(C) 4;　　(D) 5.

解 选(D)，因为

$$P\{1/2\leqslant X\leqslant 5/2\}=P\{X=2\}+P\{X=1\}=1/5,$$

$$P\{1\leqslant X\leqslant 2\}=P\{X=2\}+P\{X=1\}=1/5,$$

$$P\{0<X<3\}=P\{X=2\}+P\{X=1\}=1/5,$$

$$P\{X=1 \text{ 和 } X=2\}=P\{X=2\}+P\{X=1\}=1/5,$$

$$P\{X=3\}=3/15=1/5.$$

例10 若 $f(k)=C\lambda^k/k!$，$k=0,1,\cdots,\lambda>0$ 是离散型随机变量

的概率函数，则$C=($　　$)$.

(A) e^{λ};　　(B) $1-e^{\lambda}$;　　(C) $e^{-\lambda}$;　　(D) $1-e^{-\lambda}$.

解　选(C). 因为$\sum_{k=0}^{\infty}C\lambda^k/k!\ =Ce^{\lambda}$，故$C=e^{-\lambda}$.

例11　已知在5重伯努利试验中成功的次数不服从0-1分布，且$P\{X=1\}=P\{X=2\}$，求概率$P\{X=4\}$.

解　设在每次试验中A成功的概率为p，则

$$C_5^1p(1-p)^4=C_5^2p^2(1-p)^3\Longrightarrow p=1/3,$$

所以

$$P\{X=4\}=C_5^4\times(1/3)^4\times2/3=10/243.$$

例12　设1 h内进入某图书馆的读者人数服从泊松分布，已知1 h内无人进入图书馆的概率为0.01，求1 h内至少有两人进入图书馆的概率.

解　已知$X\sim\pi(\lambda)$，但λ未知，需由已知概率求出λ. 因为

$$P\{X=0\}=e^{-\lambda}=0.01,\quad 知\quad \lambda=2\ln10.$$

所以

$$\begin{aligned}P\{X\geqslant2\}&=1-P\{X=0\}-P\{X=1\}\\&=1-0.01-0.01\times2\ln10=0.944.\end{aligned}$$

例13　某处有5个公用电话亭，调查结果显示，在任一时刻t，每门电话被使用的概率为0.1，求在同一时刻

(1) 恰有2门电话被使用的概率；

(2) 至少有3门电话被使用的概率；

(3) 至多有3门电话被使用的概率；

(4) 没有一门电话被使用的概率.

解　X可取值为$0,1,\cdots,5$，$X\sim B(5,0.1)$，故

(1) $P\{X=2\}=C_5^2\times0.1^2\times0.9^3=0.073$.

(2) $P\{X\geqslant3\}=C_5^3\times0.1^3\times0.9^2+C_5^4\times0.1^4\times0.9+C_5^5\times0.1^5=0.0086$.

(3) $P\{X\leqslant3\}=1\times[C_5^4\times0.1^4\times0.9+C_5^5\times0.1^5]=0.9995$.

(4) $P\{X=0\}=0.9^5=0.5905$.

例 14 某家电维修站保修本地区某品牌的600台电视机，已知每台电视机的故障率为0.005.

（1）如果维修站有4名维修工，每台只需1人维修，求电视机能及时维修的概率；

（2）维修站需配备多少名维修工，才能使及时维修的概率不小于0.96？

解 设同一时刻发生故障的电视机台数为 X，$X \sim B(600,0.005)$，由于n很大，而p较小，可以利用泊松定理计算. 因为$\lambda=np=3$，所以

（1）$P\{X\leqslant 4\}=1-0.1847=0.8153$（可查表）.

（2）$P\{X\leqslant n\}\geqslant 0.96$，即$\sum_{k=n+1}^{\infty} 3^k e^{-3}/k! \leqslant 0.04$，查表知$n=6$，即需配备6名维修工.

例 15 有甲、乙、丙三支球队到某地比赛，该地只有一块训练场地，商定摸球决定哪支球队先使用场地. 摸球办法如下：盒中放两个白球、一个黑球，进行不放回的摸球，直到摸到黑球为止. 若第一次摸到黑球，则甲队先使用；第二次摸到黑球，则乙队先使用；最后一次才摸到黑球，则丙队先使用. 问：这种摸球办法公平吗？若改为放回摸球，是否公平？

解 是否公平表现为三次摸球中摸到黑球的概率是否相等的问题. 不放回摸球，摸到黑球概率是条件概率，为

$$P\{X=1\}=1/3,\quad P\{X=2\}=(1-1/3)\times 1/2=1/3,$$
$$P\{X=3\}=(1-1/3)\times(1-1/2)\times 1=1/3.$$

所以，三次摸到黑球的概率相等，是公平的.

放回摸球，第k次摸到黑球，是几何概率，为

$$P\{X=1\}=1/3,\quad P\{X=2\}=(1-1/3)\times 1/3=2/9$$
$$P\{X=3\}=(1-1/3)^2\times 1/3=4/27.$$

所以，三次摸到黑球的概率不同，是不公平的.

例 16 随机数字序列要多长才能使数字0至少出现一次的概

率不小于0.9?

解 以X记数字0出现的次数,求使$P\{X\geqslant 1\}\geqslant 0.9$的试验次数$n$.

随机数字序列是由0～9等10个数字随机取一个而排成的,取到数字0的概率为0.1.取n次即进行n次独立重复试验,所以X服从二项分布$B(n,p)$,$p=0.1$.于是由

$$P\{X\geqslant 1\}=1-P\{X=0\}=1-C_n^0\times 0.9^0\times 0.9^n\geqslant 0.9$$

得 $$0.9^n\leqslant 0.1\Longrightarrow n\lg 0.9\leqslant \lg 0.1\Longrightarrow n=22.$$

即随机数字序列至少要有22位数字,才能使数字0至少出现一次的概率不小于0.9.

至此可以看到,求离散型随机变量的事件的概率,有以下一些常用方法:

(1) 利用古典型概率、条件概率、概率运算性质与公式求事件的概率$P\{X=k\}$.

(2) 当已知X服从某种分布时,按分布律及运算性质求事件的概率$P\{X=k\}$.

(3) 当已知X的分布函数时,按分布函数定义得

$$F(k)-F(k-1)=P\{X=k\}.$$

(4) 利用概率的基本性质:$\sum p_k=1$,$p_k\geqslant 0$,建立关于p_k的方程组,解得$P\{X=k\}$.

第三节 连续型随机变量及其概率分布

主要内容

1. 概率密度函数的定义

如果对于随机变量X的分布函数$F(x)$,存在非负可积函数

$f(x)$，使对于任意实数 x，有

$$F(x)=\int_{-\infty}^{x} f(t)\mathrm{d}t,$$

则称 X 为连续型随机变量. $f(x)$称为 X 的概率密度函数.

2. 概率密度函数的性质

(1) $f(x)\geqslant 0$；

(2) $\int_{-\infty}^{+\infty} f(x)\mathrm{d}x=1$；

(3) $P\{x_1<X\leqslant x_2\}=F(x_2)-F(x_1)=\int_{x_1}^{x_2} f(x)\mathrm{d}x$；

(4) 在 $f(x)$的连续点 x 处，$F'(x)=f(x)$.

3. 常用的重要的连续型随机变量及其分布

(1) 均匀分布　记为 $X\sim U(a,b)$，概率密度为

$$f(x)=\begin{cases} 1/(b-a), & x\in(a,b), \\ 0, & \text{其它}. \end{cases}$$

在(a,b)上服从均匀分布的随机变量 X，它落在(a,b)的任一子区间内的概率只与子区间的长度有关，而与子区间的位置无关，即

$$P\{c<X\leqslant c+l\}=\int_{c}^{c+l} f(x)\mathrm{d}x=\frac{l}{b-a},$$

其中 $a\leqslant c<c+l\leqslant b$.

(2) 正态分布　记为 $X\sim N(\mu,\sigma^2)$，概率密度为

$$f(x)=\frac{1}{\sqrt{2\pi}\sigma}\mathrm{e}^{-(x-\mu)^2/2},\quad -\infty<x<+\infty,$$

其中 μ,σ^2 $(\sigma>0)$为常数. 称 X 服从参数为 μ,σ^2 的正态分布.

当 $\mu=0,\sigma^2=1$ 时，称 $X\sim N(0,1)$为标准正态分布. 若 $X\sim N(\mu,\sigma^2)$，则 $Y=(X-\mu)/\sigma\sim N(0,1)$. 标准正态分布的分布函数 $\Phi(x)$有表可查，对一般正态分布 $X\sim N(\mu,\sigma^2)$，其分布函数 $F(x)=\Phi\left(\frac{x-\mu}{\sigma}\right)$.

(3) 指数分布　记为 $X\sim \mathrm{e}(\lambda)$，其概率密度为

$$f(x)=\begin{cases} \lambda\mathrm{e}^{-\lambda x}, & x\geqslant 0, \\ 0, & x<0 \end{cases}\quad (\lambda>0,\text{常数}).$$

指数分布有一个特殊的性质：

$$P\{X>s+t|X>s\}=P\{X>t\}.$$

若以 X 表示产品的使用寿命，则上式的意义为：产品使用了时间 s 后再使用时间 t 以上的概率，等于新产品使用时间 t 以上的概率. 这种性质称为“无记忆性”，即从某时刻起产品的使用寿命与已使用时间无关.

疑 难 解 析

1. 为什么要区别连续型随机变量与离散型随机变量？

答 要了解一个随机变量 X，知道了分布函数就了解了它的统计规律性. 但 $F(x)$ 的形式和性质不尽一致，所以要分别考察.

离散型随机变量 X 的分布函数不是连续的，因此不能作求导和积分运算. 但只要了解了它的概率分布 $P\{X=k\}=p_k$，则 $F(x)$ 与事件的概率均可求得. 所以，对离散型随机变量只需求出概率分布就可以了.

连续型随机变量 X 的分布函数是连续的，$F(x)$ 可以表示为 $\int_{-\infty}^{x} f(t)\mathrm{d}t$（但 $F(x)$ 连续，不能得出 X 是连续型随机变量）. 因此了解了 X 的概率密度函数 $f(t)$，则 $F(x)$ 与事件的概率均可求得，所以，对连续型随机变量只需求出概率密度函数就可以了.

2. 连续型随机变量的 $f(t)\mathrm{d}t$ 与离散随机变量的 p_k 在概率中的意义是否相同？为什么？

答 相同. 因为，对于离散型随机变量 X 来说，$P\{X=x_k\}=p_k$ 表示 X 取 x_k 时的概率；而对连续型随机变量而言，任一点的概率为零，因此，由定积分中值定理有

$$P\{x<X\leqslant x+\Delta x\}=\int_{x}^{x+\Delta x} f(t)\mathrm{d}t\approx f(x)\mathrm{d}x.$$

在对连续型随机变量进行离散化处理的思想下，两者的概率意义

是相同的.

3. 为什么凭 $P\{X=x_k\}=0$ 不能说 $X=x_k$ 一定是不可能事件?

答 对于离散型随机变量来说, $P\{X=x_k\}=0$ 的点 $X=x_k$ 的确是不可能事件,但是对于连续型随机变量来说,任一点的概率都是零. 这由 $P\{x<X\leqslant x+\Delta x\}\approx f(x)\mathrm{d}x$ 可以看出,当 $\mathrm{d}x\to 0$ 时概率趋于零. 因此,不能只凭 $P\{X=x_k\}=0$ 就断言 $X=x_k$ 一定是不可能事件.

4. 试描述正态分布的性态.

答 正态分布与二项分布、泊松分布是概率论的三大重要分布,在实践中有广泛的应用. 一般地,如果影响某一数量指标的因素有多个,而每个随机因素的作用都不是主要的,则该数量指标必服从正态分布.

正态分布的概率密度曲线有如下性质:

(1) 曲线关于直线 $x=\mu$ 对称,在 $x=\mu$ 处取得最大值 $f(\mu)=1/(\sqrt{2\pi}\sigma)$,因此随机变量在 $x=\mu$ 附近取值的概率最大. 显然,对长度相同的区间,当区间离 μ 越远, X 落在该区间内概率越小. 曲线是单峰曲线,在 $x=\pm(\sigma+\mu)$ 处有拐点,并以 x 轴为渐近线.

(2) 固定 σ,改变 μ 值,则曲线图形不变,对称轴平移,所以 μ 又称位置参数;固定 μ,改变 σ 值,则最大值 $f(\mu)=1/(\sqrt{2\pi}\sigma)$ 改变, σ 变小时,图形变尖,因而 X 落在 μ 附近的概率也变大. 因此, σ 又表征 X 值的集中程度,称为精度参数.

对于任何 $X\sim N(\mu,\sigma^2)$,有

$$P\{a<X\leqslant b\}=\Phi\left(\frac{x-b}{\sigma}\right)-\Phi\left(\frac{x-a}{\sigma}\right),$$

且

$$\Phi(-x)=1-\Phi(x).$$

方法、技巧与典型例题分析

这一部分有三大类题:一是已知分布函数,求概率密度与事件的概率;二是已知概率密度,求事件的概率;三是已知分布形式和

概率，确定参数或x的数值.

对于这三大类题，首先要熟知分布函数与密度函数之间的关系，用分布函数与密度函数的性质解题；在已知分布与概率密度时，要善于利用分布的特点解题.

例1 已知

$$F(x)=\begin{cases}0, & x<0,\\ x+1/2, & 0\leqslant x<1/2,\\ 1, & x\geqslant 1/2,\end{cases}$$

则$F(x)$是(　　)随机变量的分布函数.

(A) 连续型；　　(B) 离散型；

(C) 非连续型；　　(D) 非连续亦非离散型.

解 选(D). 因为$F(x)$在$(-\infty,+\infty)$上单调不减，右连续，且$\lim\limits_{x\to-\infty}F(x)=0$, $\lim\limits_{x\to+\infty}F(x)=1$，所以它是一个分布函数. 又$F(x)$除$x=0,1/2$外处处可导，而

$$F'(x)=\begin{cases}0, & x<0,x>1/2,\\ 1, & \text{其它}.\end{cases}$$

但
$$\int_{-\infty}^{x}F'(x)\mathrm{d}x=\begin{cases}0, & x<0,\\ x, & 0<x\leqslant 1/2,\\ 1/2, & x>1/2,\end{cases}$$

它不等于$F(x)$，因此不存在密度函数. 又$F(x)$也不是阶跃函数，所以$F(x)$是既非离散型又非连续型的随机变量的分布函数.

例2 已知随机变量X的密度函数为

$$f(x)=A\mathrm{e}^{-|x|},\quad -\infty<x<+\infty.$$

求：(1) A值；　(2) $P\{0<X<1\}$；　(3) $F(x)$.

解 由$\int_{-\infty}^{+\infty}A\mathrm{e}^{-|x|}\mathrm{d}x=2A\int_{0}^{+\infty}\mathrm{e}^{-x}\mathrm{d}x=1$，得$2A=1$，解得$A=1/2$. 所以

$$P\{0<X<1\}=\int_0^1\frac{1}{2}\mathrm{e}^{-x}\mathrm{d}x=\frac{1}{2}(1-\mathrm{e}^{-1}).$$

当 $x<0$ 时，$$F(x)=\frac{1}{2}\int_{-\infty}^{x}e^{t}dt=\frac{1}{2}e^{x},$$

当 $x\geqslant 0$ 时，$$F(x)=\frac{1}{2}\int_{-\infty}^{0}e^{x}dx+\frac{1}{2}\int_{0}^{x}e^{-x}dx=1-\frac{1}{2}e^{-x}.$$

所以 $$F(x)=\begin{cases}e^{x}/2, & x<0,\\ 1-e^{-x}/2, & x\geqslant 0.\end{cases}$$

例3 设某种仪器内装有三只同样的电子管，电子管使用寿命 X 的概率密度函数为

$$f(x)=\begin{cases}100/x^{2}, & x\geqslant 100,\\ 0, & x<100.\end{cases}$$

求：(1) 在开始150 h内没有电子管损坏的概率；

(2) 在这段时间内有一只电子管损坏的概率；

(3) $F(x)$.

解 (1) $P\{X\leqslant 150\}=\int_{100}^{150}\frac{100}{x^{2}}dx=\frac{1}{3}$，故

$$p=[P\{X>150\}]^{3}=(1-1/3)^{3}=8/27.$$

(2) 将观察三只电子管看作三次独立重复试验，由二项概率得

$$p=C_{3}^{1}\times 1/3\times 2/3^{2}=4/9.$$

(3) 当 $x<100$ 时，$F(x)=0$；当 $x\geqslant 100$ 时，

$$F(x)=-\int_{100}^{x}\frac{100}{t^{2}}dt=1-\frac{100}{x}.$$

所以 $$F(x)=\begin{cases}0, & x<100,\\ 1-\frac{100}{x}, & x\geqslant 100.\end{cases}$$

例4 随机变量 X 的密度函数如图2.1所示，试求：

(1) 密度函数 $f(x)$；

(2) 分布函数 $F(x)$；

(3) 概率 $P\{0.2<X\leqslant 1.2\}$.

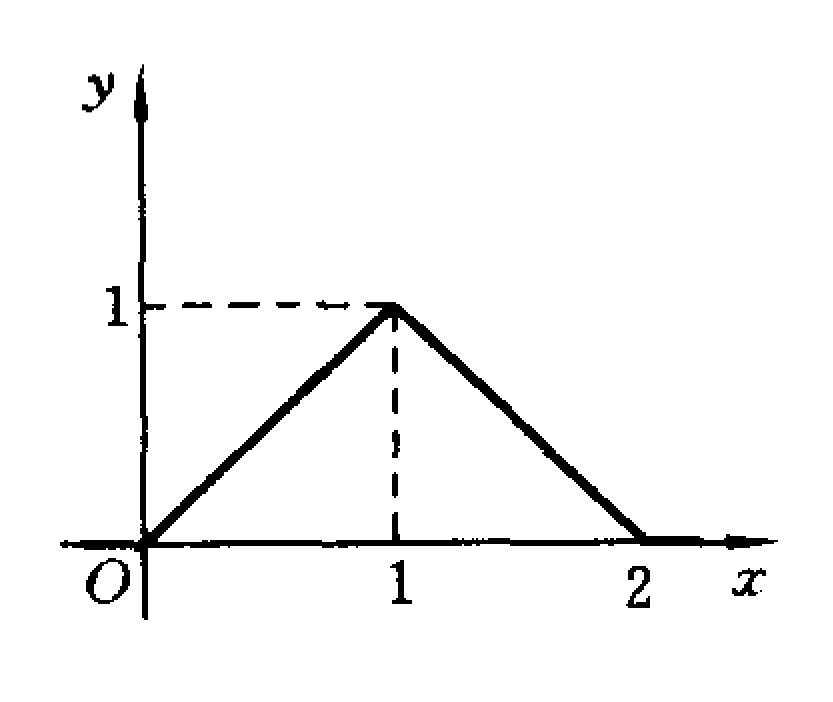


图2.1

解 (1) 由图2.1可直接写出密度函数

$$f(x)=\begin{cases}0, & x<0,\\ x, & 0\leqslant x<1,\\ 2-x, & 1\leqslant x<2,\\ 0, & x\geqslant 2.\end{cases}$$

(2) 由 $F(x)=\int_{-\infty}^{x}f(t)\mathrm{d}t$,得

$$F(x)=\begin{cases}0, & x<0,\\ x^2/2, & 0\leqslant x<1,\\ 2x-x^2/2-1, & 1\leqslant x<2,\\ 1, & x\geqslant 2.\end{cases}$$

(3) $P\{0.2<X\leqslant 1.2\}=F(1.2)-F(0.2)=0.66$.

例5 设$F(x)$是随机变量的分布函数,证明:对任何$h\neq 0$,函数

$$\Phi(x)=\frac{1}{h}\int_{x}^{x+h}F(t)\mathrm{d}t,\quad \Psi(x)=\frac{1}{2h}\int_{x-h}^{x+h}F(t)\mathrm{d}t$$

也是随机变量的分布函数.

证 只需证$\Phi(x)$,$\Psi(x)$满足分布函数的性质.

由于$F(x)$单调不减,所以对任意$\delta>0$,有

$$F(x)\leqslant F(x+\delta),\quad x\in(-\infty,+\infty),$$

使得
$$\Phi(x+\delta)=\frac{1}{h}\int_{x+\delta}^{x+\delta+h}F(t)\mathrm{d}t=\frac{1}{h}\int_{x}^{x+h}F(t+\delta)\mathrm{d}t$$
$$\geqslant\frac{1}{h}\int_{x}^{x+h}F(t)\mathrm{d}t=\Phi(x),$$

从而知$\Phi(x)$也是单调不减函数.

又$F(x)$单调有界,从而对$F(x)$的积分必为积分上限的连续函数,所以$\Phi(x)$右连续.

利用关系式,当$t\in[x,x+h]$时,

$$F(x)=\frac{1}{h}\int_{x}^{x+h}F(t)\mathrm{d}t\leqslant\frac{1}{h}\int_{x}^{x+h}F(t)\mathrm{d}t$$
$$\leqslant\frac{1}{h}\int_{x}^{x+h}F(x+h)\mathrm{d}t=F(x+h)\leqslant 1,$$

得
$$0\leqslant F(x)\leqslant\Phi(x)\leqslant F(x+h)\leqslant 1.$$

令 $x\to-\infty$，则

$$0\leqslant\lim_{x\to-\infty}\Phi(x)\leqslant\lim_{x\to-\infty}F(x+h)=0,\quad 即\quad \lim_{x\to-\infty}\Phi(x)=0;$$

令 $x\to+\infty$，则

$$1=\lim_{x\to+\infty}F(x)\leqslant\lim_{x\to+\infty}\Phi(x)\leqslant 1,\quad 即\quad \lim_{x\to+\infty}\Phi(x)=1.$$

综上所述，$\Phi(x)$满足分布函数的三条性质，故是分布函数.

同理可证，$\Psi(x)$也是分布函数.

例6 若在图2.2所示三角形ABC中任取一点P，令点P到边AB的距离为X，求X的分布函数与密度函数.

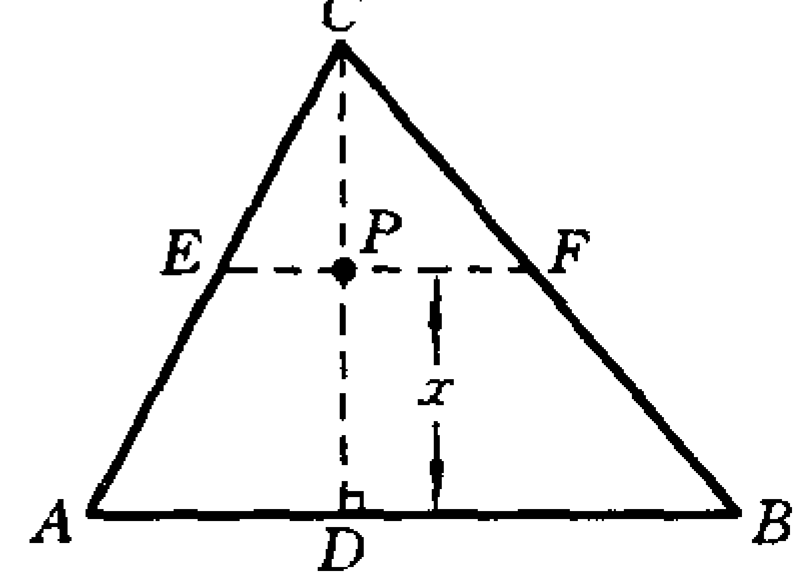


图2.2

解 设CD为AB边的高，过点P作$EF\parallel AB$，则AB与EF间距离为x. 因此，当$0\leqslant x<h$时，

$$\begin{aligned}F(x)&=P\{X\leqslant x\}\\&=\frac{梯形EFBA面积}{三角形ABC面积}\\&=1-\frac{三角形CEF面积}{三角形ABC面积}\\&=1-(h-x)^2/h^2,\end{aligned}$$

所以
$$F(x)=\begin{cases}0, & x<0,\\1-(h-x)^2/h^2, & 0\leqslant x<h,\\1, & x\geqslant h,\end{cases}$$

$$f(x)=F'(x)=\begin{cases}2(h-x)/h^2, & 0\leqslant x<h,\\0, & 其它.\end{cases}$$

例7 设某种食品的保质期(单位:d)X的概率密度函数为

$$f(x)=\begin{cases}60000/(x+100)^3, & x>0,\\0, & 其它,\end{cases}$$

求:(1)$F(x)$; (2)至少有100 d保质期的概率.

解 (1) 当$x<0$时，$F(x)=0$；当$x\geqslant 0$时，

$$F(x)=\int_0^x\frac{60000}{(t+100)^3}\mathrm{d}t=-\frac{30000}{(t+100)^2}\bigg|_0^x=1-\frac{30000}{(x+100)^2}.$$

所以 $$F(x)=\begin{cases}1-30000/(x+100)^2, & x\geqslant 0,\\ 0, & x<0.\end{cases}$$

(2) $$P\{X\geqslant 100\}=1-P\{X<100\}=1-F(100)$$
$$=30000/(100+100)^2=3/4.$$

例 8 设随机变量 X 的绝对值不大于 1，$P\{X=-1\}=1/8$，$P\{X=1\}=1/4$；在事件$\{-1<X<1\}$出现的条件下，X 在$(-1,1)$内的任一子区间上取值的条件概率与该子区间的长度成正比. 试求 X 的分布函数 $F(x)$.

解 显然，当 $x<-1$ 时，$F(x)=0$；当 $x\geqslant 1$ 时，$F(x)=1$. 所以
$$P\{-1<X<1\}=1-1/4-1/8=5/8.$$
因为 X 在$(-1,1)$内服从均匀分布，所以，当 $-1<x<1$ 时，
$$P\{-1<X\leqslant x\mid -1<X<1\}=(x+1)/2.$$
$$\begin{aligned}&P\{-1<X\leqslant x\}\\&=P\{-1<X\leqslant x,-1<X<1\}\\&=P\{-1<X<1\}P\{-1<X\leqslant x\mid -1<X<1\}\\&=5/8\times(x+1)/2=5(x+1)/16,\end{aligned}$$
即当 $-1\leqslant x<1$ 时，
$$\begin{aligned}F(x)&=P\{X=-1\}+P\{-1<X\leqslant x\}\\&=1/8+(x+1)/16=(5x+7)/16,\end{aligned}$$
故 $$F(x)=\begin{cases}0, & x<-1,\\ (5x+7)/16, & -1\leqslant x<1,\\ 1, & x\geqslant 1.\end{cases}$$

例 9 随机变量 X 的概率密度函数为
$$f(x)=\begin{cases}A\cos x, & |x|\leqslant\pi/2,\\ 0, & \text{其它},\end{cases}$$
求：(1) A 的值； (2) $F(x)$； (3) $P\{0<X\leqslant\pi/4\}$.

解 (1) 由 $1=\int_{-\pi/2}^{\pi/2}A\cos x\,\mathrm{d}x=2A\int_0^{\pi/2}\cos x\,\mathrm{d}x=2A$，得 $A=1/2$.

(2) 当 $x<-\pi/2$ 时，$F(x)=0$；当 $x\geqslant\pi/2$ 时，$F(x)=1$；当

$-\pi/2 \leqslant x < \pi/2$ 时，

$$F(x)=\int_{-\pi/2}^{x}\frac{1}{2}\cos x\mathrm{d}x=\frac{1}{2}+\frac{1}{2}\sin x.$$

所以 $$F(x)=\begin{cases}0, & x<-\pi/2,\\ 1/2+(\sin x)/2, & \pi/2\leqslant x<\pi/2,\\ 1, & x\geqslant \pi/2.\end{cases}$$

(3) $P\{0<X\leqslant \pi/4\}=F(\pi/4)-F(0)=\sqrt{2}/4.$

例 10 设随机变量 X 的分布函数是

$$F(x)=\begin{cases}A+B\mathrm{e}^{-x^2/2}, & x>0,\\ 0, & x\leqslant 0,\end{cases}$$

求：(1) A 与 B 的值； (2) $f(x)$； (3) $P\{1<X<2\}$.

解 (1) 由 $0=F(0)=A+B$，$1=F(+\infty)=A$，得

$$A=1,\quad B=-1.$$

(2) $$f(x)=F'(x)=\begin{cases}x\mathrm{e}^{-x^2/2}, & x>0,\\ 0, & x\leqslant 0.\end{cases}$$

(3) $P\{1<X<2\}=F(2)-F(1)=0.47.$

例 11 设随机变量 X 的分布函数为

$$F(x)=\begin{cases}0, & x<1,\\ \ln x, & 1\leqslant x<\mathrm{e},\\ 1, & x\geqslant \mathrm{e},\end{cases}$$

求：(1) $P\{X<2\}$，$P\{0<X\leqslant 3\}$，$P\{2<X<5/2\}$；

(2) 概率密度函数 $f(x)$.

解 (1) 利用分布函数求概率，即

$$P\{X<2\}=F(2)=\ln 2,$$

$$P\{0<X\leqslant 3\}=F(3)-F(0)=1-0=1,$$

$$P\{2<X<5/2\}=F(5/2)-F(2)=\ln(5/2)-\ln 2=\ln(5/4).$$

(2) 由 $f(x)=F'(x)$ 可得

$$f(x)=\begin{cases}0, & x<1 \text{ 或 } x>\mathrm{e},\\ 1/x, & 1\leqslant x<\mathrm{e}.\end{cases}$$

例 12　设在区间$[a,b]$上，随机变量 X 的密度函数为 $f(x)=\sin x$，而在$[a,b]$外，$f(x)=0$，则区间$[a,b]$等于(　　).

(A) $[0,\pi/2]$；　　　　(B) $[0,\pi]$；

(C) $[-\pi/2,0]$；　　　　(D) $[0,3\pi/2]$.

解　选(A). 因为在$[0,\pi/2]$上，$\sin x\geqslant 0$(非负可积)，而且 $\int_0^{\pi/2}\sin x\mathrm{d}x=1$，故 $f(x)$是概率密度函数.

在$[0,\pi]$上，$\int_0^{\pi}\sin x\mathrm{d}x=2\neq 1$，所以 $f(x)$不是概率密度函数.

在$[-\pi/2,0]$上，$\sin x\leqslant 0$，所以 $f(x)$不是概率密度函数.

在$[0,3\pi/2]$上，当 $x>\pi$ 时，$\sin x<0$，所以 $f(x)$不是概率密度函数.

例 13　设连续型随机变量 X 的密度函数是

$$f(x)=\begin{cases}\dfrac{1}{C}x\mathrm{e}^{-x^2/(2C)}, & x>0,\\ 0, & \text{其它},\end{cases}$$

则式中 C 为(　　).

(A) 任何实数；　　　　(B) 正数；

(C) 1；　　　　(D) 任何非零实数.

解　选(B)、(C). 由密度函数性质，

$$\int_0^{+\infty}\frac{1}{C}x\mathrm{e}^{-x^2/(2C)}\mathrm{d}x=\int_0^{+\infty}\mathrm{e}^{-x^2/(2C)}\mathrm{d}\,\frac{x^2}{2C}$$

$$=-\mathrm{e}^{-x^2/(2C)}\Big|_0^{+\infty}=\mathrm{e}^{-\frac{0}{2C}}=1,$$

当 $C>0$ 时，即有上述等式成立.

例 14　设随机变量 X 在$[2,5]$上服从均匀分布. 现对 X 进行三次独立观察，求至少有两次的观察值大于 3 的概率.

解　因为 $X\sim U(2,5)$，所以

$$f(x)=\begin{cases}1/3, & x\in[2,5],\\ 0, & \text{其它}.\end{cases}$$

设 $A=P\{X>3\}=\int_3^5 \frac{1}{3}\mathrm{d}x=\frac{2}{3}$，依二项概率公式，

$$p=C_3^2\times(2/3)^2\times(1-2/3)+C_3^3\times(2/3)^3=20/27.$$

例 15 设随机变量 X 的概率密度函数

$$f(x)=\begin{cases}e^{-(x-a)}, & x>x_0,\\ 0, & 其它,\end{cases}$$

则 $x_0=$ ________.

解 因为

$$1=\int_{x_0}^{+\infty}e^{-(x-a)}\mathrm{d}x=-e^{-(x-a)}\Big|_{x_0}^{+\infty}=e^{-(x_0-a)}=1,$$

所以 $x_0-a=0$，即 $x_0=a$.

例 16 设某河流每年的最高洪水水位（单位：m）X 的概率密度为

$$f(x)=\begin{cases}2/x^3, & x\geqslant 1,\\ 0, & x<1.\end{cases}$$

今要修建能防御百年一遇洪水（即 $p\leqslant 0.01$）的河堤，问：河堤应修多高？（河堤高度起点与洪水水位起点相同.）

解 设河堤高为 h，则应有 $P\{X\geqslant h\}=0.01$，由

$$P\{X\geqslant h\}=1-\int_1^h \frac{2}{x^3}\mathrm{d}x=\frac{1}{h^2}=0.01\Longrightarrow h=10,$$

所以，河堤应修 10 m 高.

例 17 某种螺栓的长度（单位：cm）$X\sim N(10.05,0.06^2)$. 若规定长度在范围 10.05 ± 0.12 内为合格品，求任取一螺栓为不合格品的概率.

解 $X\sim N(10.05,0.06^2)$，则

$$(X-10.05)/0.06\sim N(0,1).$$

$$P\{(10.05-0.12)\leqslant X\leqslant(10.05+0.12)\}$$
$$=\Phi[(10.17-10.05)/0.06]-\Phi[(9.93-10.05)/0.06]$$
$$=\Phi(2)-\Phi(-2)=2\Phi(2)-1=0.9544,$$

所以，任取一螺栓为不合格品的概率

$$p=1-0.9544=0.0456.$$

例 18 设一大型设备在任何长为 t 的时间内发生故障的次数 $N(t)$ 服从参数为 λt 的泊松分布，试求：

(1) 相继两次故障的时间间隔 T 的概率分布；

(2) 在设备已经无故障工作 8 h 的情况下，再无故障工作 8 h 的概率.

解 当 $T>t$ 时，即 $N(t)=0$（也就是相继两次故障的间隔时间 $T>t$，表示在时间段 t 内没发生故障），从而

(1) $F(t)=P\{T\leqslant t\}=1-P\{T>t\}=1-P\{N(t)=0\}$

$$=1-\frac{(\lambda t)^0}{0!}e^{-\lambda t}=1-e^{-\lambda t}.$$

(2) $p=P\{T\geqslant 16|T\geqslant 8\}$ 的求法有两种.

一种是：由 $F(t)=1-e^{-\lambda t}$，知 T 服从参数为 λ 的指数分布，依指数分布的无记忆性 $P\{X>s+t|X>s\}=P\{X>t\}$，得

$$p=P\{T\geqslant 16|T\geqslant 8\}=P\{T\geqslant 8\}=e^{-8\lambda}.$$

另一种是：按条件概率直接计算，得

$$\begin{aligned}p&=P\{T\geqslant 16|T\geqslant 8\}=P\{T\geqslant 16,T\geqslant 8\}/P\{T\geqslant 8\}\\&=P\{T\geqslant 16\}/P\{T\geqslant 8\}\\&=[1-P\{T<16\}]/[1-P\{T<8\}]\\&=[1-F(16)]/[1-F(8)]=e^{-16\lambda}/e^{-8\lambda}=e^{-8\lambda}.\end{aligned}$$

例 19 某人每天上班有两条线可走，第一条路线较短，但容易堵车，所需时间（单位：min）X 的概率密度为

$$f(x)=\begin{cases}\dfrac{1}{5\sqrt{2\pi}}e^{-(x-40)^2/200}, & x>40,\\ 0, & x\leqslant 40;\end{cases}$$

第二条线路较长，但不易堵车，所需时间 Y 的概率密度为

$$f(y)=\begin{cases}\dfrac{1}{2\sqrt{2\pi}}e^{-(x-50)^2/32}, & x>50,\\ 0, & x\leqslant 50.\end{cases}$$

问：(1)如果提前 60 min 离家，走哪条路线上班迟到的可能性小？

(2)如果只能提前 55 min 离家，走哪条路线上班迟到的可能性小？

解 计算超过给定时间的概率，选择概率小的路线．由于 $f(x)$ 和 $f(y)$ 都不是正态分布的概率密度函数，在计算中可以设法化为正态分布来求，使计算更简单．

(1) 因为，当时间超过 60 min 时，

$$P\{X>60\}=\frac{1}{5}\int_{60}^{+\infty}\frac{1}{\sqrt{2\pi}}e^{-(x-40)^2/200}dx \quad \left(令\frac{x-40}{10}=t\right)$$

$$=\frac{2}{\sqrt{2\pi}}\int_{2}^{+\infty}e^{-t^2/2}dt=2[1-\Phi(2)]=0.0456,$$

$$P\{Y>60\}=\frac{1}{2\sqrt{2\pi}}\int_{60}^{+\infty}e^{-(x-50)^2/32}dx \quad \left(令\frac{x-50}{4}=t\right)$$

$$=\frac{2}{\sqrt{2\pi}}\int_{2.5}^{+\infty}e^{-t^2/2}dt=2[1-\Phi(2.5)]=0.0124.$$

可见 $P\{X>60\}>P\{Y>60\}$，所以选择第二条路线．

(2) 因为，当时间超过 55 min 时，

$$P\{X>55\}=\frac{2}{\sqrt{2\pi}}\int_{1.5}^{+\infty}e^{-t^2/2}dt=2[1-\Phi(1.5)]=0.1336,$$

$$P\{Y>55\}=\frac{2}{\sqrt{2\pi}}\int_{1.25}^{+\infty}e^{-t^2/2}dt=2[1-\Phi(1.25)]=0.2112.$$

可见 $P\{X>55\}<P\{Y>55\}$，所以选择第一条路线．

例 20 设随机变量 $X\sim N(0,\sigma^2)$，问：当 σ 取何值时，X 落入区间(1,3)的概率最大？

解 因为 $X\sim N(0,\sigma^2)$，所以 $F(x)=\Phi\left(\frac{x}{\sigma}\right)$，于是

$$F(1<x<3)=\Phi\left(\frac{3}{\sigma}\right)-\Phi\left(\frac{1}{\sigma}\right).$$

令上式等于 $g(\sigma)$，利用微积分中求极值的方法，有

$$g'(\sigma)=\Phi'\left(\frac{3}{\sigma}\right)\left(\frac{-3}{\sigma^2}\right)+\Phi'\left(\frac{1}{\sigma}\right)\frac{1}{\sigma^2}$$

$$=\frac{1}{\sqrt{2\pi}\sigma^3}e^{-1/(2\sigma^2)}[1-3e^{-8/(2\sigma^2)}].$$

令上式等于零，解得$\sigma_0^2=4/\ln 3$. 又$g''(\sigma_0)<0$，故$\sigma=2/\sqrt{\ln 3}$为极大值点，且唯一. 所以，当$\sigma=2/\sqrt{\ln 3}$时，X 落入区间$(1,3)$的概率最大.

例 21 设随机变量$X\sim N(\mu,\sigma^2)$，求概率$P\{|X-\mu|<k\sigma\}$，$k=1,2,3$.

解 此即著名的"三σ规则".

$$\begin{aligned}P\{|X-\mu|<k\sigma\}&=P\{\mu-k\sigma<X<\mu+k\sigma\}\\&=\Phi\left(\frac{\mu-k\sigma-\mu}{\sigma}\right)-\Phi\left(\frac{\mu+k\sigma-\mu}{\sigma}\right)\\&=\Phi(k)-\Phi(-k)=2\Phi(k)-1.\end{aligned}$$

当$\sigma=1$时， $P\{|X-\mu|<\sigma\}=2\Phi(1)-1=0.6826$；

当$\sigma=2$时， $P\{|X-\mu|<2\sigma\}=2\Phi(2)-1=0.9544$；

当$\sigma=3$时， $P\{|X-\mu|<3\sigma\}=2\Phi(3)-1=0.9974$.

例22 设$X\sim N(60,9)$，求使X落入区间$(-\infty,x_1]$，$(x_1,x_2]$，$(x_2,+\infty]$的概率比为3∶4∶5的分点x_1,x_2.

解 由$X\sim N(60,9)$知，$(X-60)/3\sim N(0,1)$，故

$$P\{X\leqslant x_1\}=\Phi\left(\frac{x_1-60}{3}\right)=\frac{3}{3+4+5}=0.25,$$

即 $$1-\Phi\left(\frac{60-x_1}{3}\right)=0.25\Longrightarrow\Phi\left(\frac{60-x_1}{3}\right)=0.75.$$

查表知$(60-x_1)/3=0.075$，故$x_1=57.975$.

$$P\{X>x_2\}=1-\Phi\left(\frac{x_2-60}{3}\right)=\frac{5}{3+4+5}=0.4167,$$

即 $$\Phi\left(\frac{x_2-60}{3}\right)=0.5833.$$

查表知$(x_2-60)/3=2.1$，故$x_2=60.63$.

例 23 设$X\sim N(10,\sigma^2)$，且$P\{10<X<20\}=0.3$，求

$$P\{0<X<10\}.$$

解 由正态分布的性态知，X 的概率分布关于直线 $x=\mu=10$ 对称，所以

$$P\{0<X<10\}=P\{10<X<20\}.$$

例 24 设 $X\sim N(10,4)$，求：

(1) $P\{7<X<15\}$； (2) 确定 d，使 $P\{|X-10|<d\}=0.8$.

解 由 $X\sim N(10,4)$ 知，$(X-10)/2\sim N(0,1)$，故

(1) $$P\{7<X<15\}=\Phi\left(\frac{15-10}{2}\right)-\Phi\left(\frac{7-10}{2}\right)$$
$$=\Phi\left(\frac{5}{2}\right)-\Phi\left(-\frac{3}{2}\right)$$
$$=\Phi(2.5)+\Phi(1.5)-1=0.927\text{(查表)}.$$

(2) $$P\{|X-10|<d\}$$
$$=\Phi\left(\frac{10-10+d}{2}\right)-\Phi\left(\frac{10-10-d}{2}\right)$$
$$=\Phi\left(\frac{d}{2}\right)-\Phi\left(-\frac{d}{2}\right)=2\Phi\left(\frac{d}{2}\right)-1=0.8.$$

从而 $\Phi\left(\frac{d}{2}\right)=0.9$，查表知 $\frac{d}{2}=1.281$，所以 $d=2.562$.

例 25 设 X 的概率密度函数为

$$f(x)=\frac{1}{\sqrt{6\pi}}e^{-(x^2+4x-4)/6},\ -\infty<x<+\infty.$$

求：(1) $P\{1<x<3\}$； (2) 使 $\int_C^{+\infty}f(x)dx=\int_{-\infty}^{C}f(x)dx$ 的 C.

解 因为 $\frac{1}{\sqrt{6\pi}}e^{-(x^2+4x-4)/6}=\frac{1}{\sqrt{2\pi}\sqrt{3}}e^{-(x-2)^2/(2\times 3)}$，所以，$X\sim N(2,3)$，从而

$$P\{1<X<3\}=\Phi\left(\frac{3-2}{\sqrt{3}}\right)-\Phi\left(\frac{1-2}{\sqrt{3}}\right)=2\Phi\left(\frac{\sqrt{3}}{3}\right)-1$$
$$=2\Phi(0.5773)-1=0.438\text{(查表)}.$$

(2) 要使 $\int_C^{+\infty}f(x)dx=\int_{-\infty}^{C}f(x)dx$，则 C 为概率分布的对称点. 由正态分布性态知 $C=\mu=2$ 为所求.

例 26 一轰炸机带着三枚炸弹投向敌方目标，若炸弹落在目标中心 40 m 内，目标将被摧毁. 设在使用瞄准器投弹时，弹着点 X 的概率密度函数为

$$f(x)=\begin{cases}(100+x)/10000, & -100<x\leqslant 0,\\ (100-x)/10000, & 0<x\leqslant 100,\\ 0, & \text{其它}.\end{cases}$$

求投三枚炸弹后，目标被炸毁的概率.

解 一枚炸弹落在目标中心 40 m 内的概率为

$$\int_{-40}^{40}f(x)\mathrm{d}x=\frac{1}{10000}\left[\int_{-40}^{0}(100+x)\mathrm{d}x+\int_{0}^{40}(100-x)\mathrm{d}x\right]$$

$$=\frac{2}{10000}\int_{0}^{40}(100-x)\mathrm{d}x=0.64,$$

则炸弹落在 40 m 外的概率为

$$p=1-0.64=0.36.$$

所以，三枚炸弹都落在目标中心 40 m 外的概率是 0.36^3. 于是，目标被炸毁的概率是

$$p=1-0.36^3=0.953.$$

例27 公共汽车车门的高度是按成年男子与门楣碰头的概率不大于 0.01 设计的. 设成年男子身高(单位:cm)$X\sim N(170,6^2)$，试确定车门应设计的最低高度 h.

解 设车门高度为 h，则应有 $P\{X>h\}\leqslant 0.01$.

$$P\{X>h\}=1-P\{X\leqslant h\}=1-\Phi\left(\frac{h-170}{6}\right)\leqslant 0.01,$$

即 $\Phi\left(\frac{h-170}{6}\right)\geqslant 0.99$. 查表知 $\frac{h-170}{6}\geqslant 2.33$，于是

$$h=170+2.33\times 6=183.98,$$

所以，车门最低高度应为 184 cm.

例 28 对一批新子弹，任意抽取 5 发试射，如果没有一颗子弹落在离开靶心 2 m 以外，则该批子弹被接受. 设弹着点与靶心的距离 X 的概率密度函数为

$$f(x)=\begin{cases}\dfrac{2xe^{-x^2}}{1-e^{-9}}, & 0<x<3,\\ 0, & 其它,\end{cases}$$

求该批子弹被接受的概率.

解 任取一颗子弹落在靶心2 m以内的概率是

$$P\{0\leqslant X<2\}=\int_0^2\frac{2xe^{-x^2}}{1-e^{-9}}dx=\frac{-e^{-x^2}}{1-e^{-9}}\bigg|_0^2=\frac{1-e^{-4}}{1-e^{-9}},$$

所以,该批子弹被接受的概率是$\left(\dfrac{1-e^{-4}}{1-e^{-9}}\right)^5$.

第四节 随机变量的函数的分布

主要内容

1. 随机变量的函数及其分布

在一些试验中,某些随机变量往往不能直接通过观察得到,但它是某个能直接观察的随机变量的函数.这些随机变量称为随机变量的函数,它们的分布称为随机变量函数的分布.一般地,随机变量的函数的分布可以借助随机变量的分布获得.

对随机变量的函数$Y=g(X)$,当X是离散型时,Y也是离散型随机变量.当X是连续型时,Y一般也是连续型随机变量,但也可以不是连续型随机变量.

2. 定理

设随机变量X有概率密度函数$f_X(x)$,$-\infty<x<+\infty$,又设$y=g(x)$处处可导,且有$g'(x)>0$或$g'(x)<0$(即$g(x)$严格单调),则$Y=g(X)$是连续型随机变量,概率密度为

$$f_Y(y)=\begin{cases}f_X[h(y)]|h'(y)|, & \alpha<y<\beta,\\ 0, & 其它,\end{cases}$$

其中$h(y)$是$g(x)$的反函数，

$$\alpha=\min\{g(-\infty),g(+\infty)\},\quad \beta=\max\{g(-\infty),g(+\infty)\}.$$

该定理可以推广到$g(x)$是分段严格单调的情形，此时

$$f_Y(y)=f_X[h_1(y)]|h_1'(y)|+f_X[h_2(y)]|h_2'(y)|+\cdots.$$

疑难解析

1. 离散型随机变量的函数为什么一定是离散型随机变量？连续型随机变量的函数为什么不一定是连续型随机变量？

答 对离散型随机变量来说，X 的可取值为有限个或可列无穷多个，因而$Y=g(X)$的可取值也为有限个或可列无穷多个（当Y有相同值可合并时，取值个数比 X 取值个数减少），故知Y 也是离散型的随机变量.

对连续型随机变量而言，$Y=g(X)$的可取值因归类合并等原因，可能只有有限个或可列无穷多个. 这时，Y 成为离散型随机变量. 有时$Y=g(X)$的分布可以既不是阶跃函数，也不是连续函数，这时$f_Y(y)$不存在，Y 也不是连续型随机变量（见例 15）.

2. 计算随机变量的函数的分布时应注意哪些问题？

答 首先应准确确定Y 的取值范围，一般地，由$y=g(x)$即决定了Y 的取值范围. 但在离散型变量的情形，要注意相同值的合并.

其次应正确计算Y 的分布，特别是连续型随机变量X 的函数的情形. 当$y=g(x)$单调或分段单调时，可按定理写出$f_Y(y)$；否则应先求出$F_Y(y)$，再求$f_Y(y)$.

方法、技巧与典型例题分析

一、离散型随机变量X 的函数$g(X)$的概率分布的求法

对于 X 的每一取值x_i，写出 Y 的对应取值$y_i=g(x_i)$，再由$P\{Y=y_i\}=P\{X=x_i\}$写出Y 的概率分布. 如果y_i 有相等的值，则

应将相等项的概率合并，得到规范的Y的概率分布.Y的概率分布一般仍用分布律的形式写出.

二、连续型随机变量X的函数$g(X)$的概率密度函数的求法

一般有两种方法.

(1) 分布函数法　先设法求出$F_Y(y)$，再求导得出$f_Y(y)$，即由$F_Y(y)=P\{Y\leqslant y\}$进行代换$Y=g(X)$，得$F_Y(y)=P\{g(x)\leqslant y\}$；由反函数$x=g^{-1}(y)$，得$F_Y(y)=P\{X\leqslant g^{-1}(y)\}=F_X(g^{-1}(y))$，求得$F_Y(y)$；然后再利用在连续点$x$，有$F'_Y(y)=f_Y(y)$，求得$f_Y(y)$.

要注意的是：第一，反函数$x=g^{-1}(y)$是否一定存在，若不存在，则此方法不可用；第二，对y的不同值，$F_Y(y)$不一定相同(即$F_Y(y)$可能是分段函数)，解题时要注意讨论.

(2) 当$y=g(x)$符合定理条件(即$y=g(x)$严格单调、可导)，则直接套用公式就能得到

$$f_Y(y)=\begin{cases} f_X(h(y))\,|h'(y)|, & \alpha<y<\beta, \\ 0, & \text{其它}. \end{cases}$$

要注意的是：第一，要正确求出$h(y)$，并确定α,β；对分段单调情形，要求出$h_1(y),h_2(y),\cdots$，并求出α,β. 第二，代入公式时要正确计算$f_X(h(y))$.

例1　设随机变量X的分布律为

$$P\{X=x_i\}=p_i \quad (i=1,2,\cdots),$$

求随机变量以下两种情况下Y的分布律：

(1)$Y=CX$；　(2)$Y=3X^2$.

解　因为X是离散型随机变量，所以Y也是离散型的.

(1) 由题给条件知，事件$\{Y=CX_i\}$与事件$\{X=x_i\}$等价，因而$Y=CX$的分布律为

$$P\{Y=Cx_i\}=P\{X=x_i\}=p_i \quad (i=1,2,\cdots).$$

(2) 此时，要考虑x_i的取值范值. 若x_i只取正值，则事件$\{Y=3x_i^2\}$与事件$\{X=x_i\}$等价，于是得出

$$P\{Y=3x_i^2\}=P\{X=x_i\}=p_i \quad (i=1,2,\cdots).$$

若 x_i 可取正、负值，则事件$\{Y=3x_i^2\}$与事件$\{X=x_i \cup X=-x_i\}$等价，因而$Y=3x^2$ 的分布律为

$$P\{Y=3x_i^2\}=P\{X=x_i\}+P\{X=-x_i\} \quad (i=1,2,\cdots).$$

例 2 设随机变量 X 的分布律如下：

X	-1	0	1	2
p_k	0.1	0.2	0.3	0.4

求$Y_1=2X+1$ 与$Y_2=X^2$ 的分布律.

解 Y_1 的取值为$-1,1,3,5$，与 X 取值个数相等，因此概率不必合并，得Y_1 的分布律如下：

$Y_1=2X+1$	-1	1	3	5
p_k	0.1	0.2	0.3	0.4

Y_2 的取值只有 0,1,4 三个，当$X=-1$ 与 1 时，$Y_2=1$，所以概率要合并，得Y_2 的分布律如下：

$Y_2=X^2$	0	1	4
p_k	0.2	0.4	0.4

例 3 设$P\{X=k\}=(1/2)^k$，$k=1,2,\cdots$，令

$$Y=\begin{cases} 1, & \text{当} X \text{取偶数时}, \\ -1, & \text{当} X \text{取奇数时}, \end{cases}$$

求随机变量 X 的函数 Y 的分布律.

解 利用概率的加法公式.

$$P\{Y=1\}=P\{X=2\}+P\{X=4\}+\cdots+P\{X=2k\}+\cdots$$

$$=\frac{1}{4}+\frac{1}{16}+\cdots=\left(\frac{1}{4}\right)\Big/\left(1-\frac{1}{4}\right)=\frac{1}{3}.$$

$$P\{Y=-1\}=1-\frac{1}{3}=\frac{2}{3}.$$

例 4 已知随机变量 X 的分布律为

X	1	2	3	$\cdots$	n	$\cdots$
p_k	1/2	$(1/2)^2$	$(1/2)^3$	$\cdots$	$(1/2)^n$	$\cdots$

求随机变量 X 的函数 $Y=\sin(\pi X/2)$ 的分布律.

解 因为

$$Y=\sin(\pi X/2)=\begin{cases}-1, & 当 X=4k-1,\\ 0, & 当 X=2k, \quad k=0,1,2,\cdots,\\ 1, & 当 X=4k-3,\end{cases}$$

所以,Y 只有 $-1,0,1$ 三个值,由等价关系得

$$P\{Y=-1\}=\left(\frac{1}{2}\right)^3+\left(\frac{1}{2}\right)^7+\cdots+\left(\frac{1}{2}\right)^{4k-1}+\cdots$$

$$=\frac{1}{8(1-1/16)}=\frac{2}{15},$$

$$P\{Y=0\}=\left(\frac{1}{2}\right)^2+\left(\frac{1}{2}\right)^4+\cdots+\left(\frac{1}{2}\right)^{4k}+\cdots=\frac{1}{4(1-1/4)}$$

$$=\frac{1}{3},$$

$$P\{Y=1\}=\frac{1}{2}+\left(\frac{1}{2}\right)^5+\cdots+\left(\frac{1}{2}\right)^{4k-3}+\cdots+\frac{1}{2(1-1/16)}$$

$$=\frac{8}{15}.$$

故 $Y=\sin(\pi X/2)$ 的分布律为

Y	-1	0	1
p_k	$2/15$	$1/3$	$8/15$

例 5 测量一类圆形物体的半径 X 为随机变量,其分布律为

X	10	11	12	13
p_k	0.1	0.4	0.3	0.2

求圆周长 Y_1 和圆面积 Y_2 的分布律.

解 $Y_1=2\pi X$ 和 $Y_2=\pi X^2$ 都是 X 的函数,Y_1 和 Y_2 各自的值均不相等,不需合并,所以 Y_1 和 Y_2 的分布律分别为

Y_1	20π	22π	24π	26π
p_k	0.1	0.4	0.3	0.2

Y_2	100π	121π	144π	169π
p_k	0.1	0.4	0.3	0.2

例 6　设通过点(0,1)的直线与 x 轴的正向交角为 θ $(0<\theta<\pi)$，求这直线在 x 轴上截距 X 的概率密度函数.

解　设直线与 x 轴正向交角 θ 是随机变量，$\theta \sim U(0,\pi)$，概率密度为

$$f_\theta(\theta)=\begin{cases}1/\pi, & 0<\theta<\pi, \\ 0, & \text{其它}.\end{cases}$$

而 $x=g(\theta)=-\cot\theta$，即 $X=g(\theta)=-\cot\theta$. 反函数 $\theta=h(x)=\operatorname{arccot}(-x)$在$(-\infty,+\infty)$内取值，且 $g(\theta)$在$(0,\pi)$内单调并处处可导，$h'(y)=-1/(1+x^2)$. 故依定理有

$$f_X(x)=f_\theta[h(x)]\,|-1/(1+x^2)|=1/[\pi(1+x^2)],$$

其中$-\infty<x<+\infty$.

例7　设随机变量$X\sim U(-\pi/2,\pi/2)$，求随机变量$Y=\sin X$ 的概率密度函数 $f_Y(y)$.

解　因为 $y=\sin x$ 在$(-\pi/2,\pi/2)$上单调、可导，$x=h(y)=\arcsin y$，$h'(y)=1/\sqrt{1-y^2}$，由于

$$f_X(x)=\begin{cases}1/\pi, & x\in(-\pi/2,\pi/2), \\ 0, & \text{其它},\end{cases}$$

$$\min_{-\frac{\pi}{2}\leqslant x\leqslant\frac{\pi}{2}}\{\sin x\}=-1,\quad \max_{-\frac{\pi}{2}\leqslant x\leqslant\frac{\pi}{2}}\{\sin x\}=1.$$

所以依定理有

$$f_Y(y)=\begin{cases}f_X[h(y)]\,|h'(y)| \\ 0\end{cases}=\begin{cases}1/(\pi\sqrt{1-y^2}), & -1<y<1, \\ 0, & \text{其它}.\end{cases}$$

例 8　已知随机变量 X 的概率密度

$$f_X(x)=\begin{cases}2x/\pi^2, & 0<x<\pi, \\ 0, & \text{其它},\end{cases}$$

求随机变量 $Y=\sin X$ 的概率密度.

解 $y=\sin x$ 在 $(0,\pi)$ 不单调，所以要用分布函数法．因为 $0<y<1$，所以，由图2.3知

$$F_Y(y)=P\{Y\leqslant y\}=P\{\sin X\leqslant y\}$$
$$=P\{(0<X\leqslant x_1)\cup(x_2\leqslant X<\pi)\},$$

其中 $x_1=\arcsin y$， $x_2=\pi-\arcsin y$.

图2.3

由 $f_Y(y)=F'_Y(y)$，得

$$f_Y(y)=\left(\int_0^{x_1}f_X(x)\mathrm{d}x+\int_{x_2}^{\pi}f_X(x)\mathrm{d}x\right)'$$
$$=f_X(x_1)\frac{\mathrm{d}x_1}{\mathrm{d}y}-f_X(x_2)\frac{\mathrm{d}x_2}{\mathrm{d}y}$$
$$=\frac{2}{\pi^2\sqrt{1-y^2}}[\arcsin y+(\pi-\arcsin y)]=\frac{2}{\pi\sqrt{1-y^2}}.$$

于是
$$f_Y(y)=\begin{cases}2/(\pi\sqrt{1-y^2}), & 0<y<1,\\ 0, & 其它.\end{cases}$$

作图作为一种方法，可以达到非常好的直观效果，可以帮助我们正确分析和计算，希望读者很好地学习和掌握作图方法.

例9 设随机变量 $X\sim N(0,1)$，求 $Y=2X^2+1$ 的概率密度函数.

解 $y=2x^2+1$ 不是单调函数，只能用分布函数法求解．当 $y<1$ 时，$F_Y(y)=0$；当 $y\geqslant1$ 时，

$$F_Y(y)=P\{Y\leqslant y\}=P\left\{X^2\leqslant\frac{y-1}{2}\right\}=P\{-x_1<X<x_1\},$$

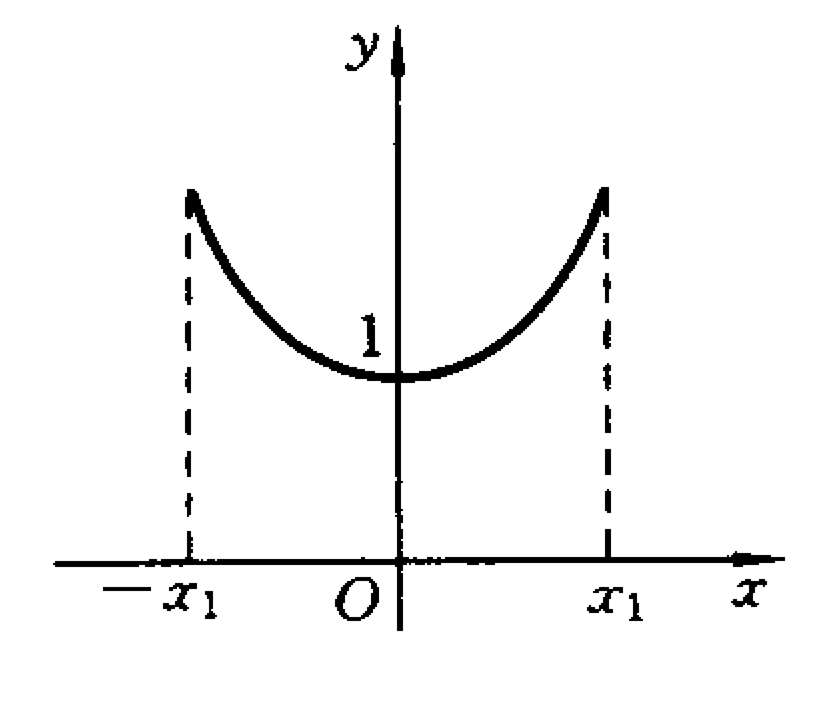


图2.4

其中 $x_1=\sqrt{(y-1)/2}$．于是，由图2.4知

$$F_Y(y)=\int_{-x_1}^{x_1}\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-x^2/2}\mathrm{d}x$$
$$=\int_{-\sqrt{(y-1)/2}}^{\sqrt{(y-1)/2}}\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-x^2/2}\mathrm{d}x$$
$$=2\int_0^{\sqrt{(y-1)/2}}\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-x^2/2}\mathrm{d}x.$$

故 $$f_Y(y)=F'_Y(y)=\frac{1}{2\sqrt{\pi(y-1)}}e^{-(y-1)/4},$$

即 $$f_Y(y)=\begin{cases}\frac{1}{2\sqrt{\pi(y-1)}}e^{-(y-1)/4}, & y\geqslant 1,\\ 0, & y<1.\end{cases}$$

例 10 已知一只昆虫所生的虫卵数 $X\sim\pi(\lambda)$，而每个虫卵发育为幼虫的概率为 p，且各虫卵是否发育为幼虫是相互独立的. 求一只昆虫所生幼虫数 Y 的概率分布.

解 因为 $X\sim\pi(\lambda)$，即 $P\{X=k\}=\lambda^k e^{-\lambda}/k!$，$k=0,1,2,\cdots$. x 个虫卵能够发育为幼虫的个数 $Y\sim B(x,p)$，所以

$$P\{Y=y|X=x\}=C_x^y p^y(1-p)^{x-y},\quad y=0,1,2,\cdots,x.$$

由全概率公式，得

$$P\{Y=y\}=\sum_{x=0}^{n}P\{X=x\}P\{Y=y|X=x\},\quad n\in\mathbf{N}.$$

而必有 $y\leqslant x$，故

$$P\{Y=y\}=\sum_{x=y}^{n}\frac{\lambda^x e^{-\lambda}}{x!}\cdot\frac{x!\ p^y q^{x-y}}{y!\ (x-y)!}=\frac{(\lambda p)^y e^{-\lambda}}{y!}\sum_{x=y}^{n}\frac{(\lambda q)^{x-y}}{(x-y)!}$$

$$=\frac{(\lambda p)^y e^{-\lambda}}{y!}\sum_{k=0}^{n-y}\frac{(\lambda q)^k}{k!}\ (q=1-p).$$

令 $n\to\infty$，即得

$$P\{Y=y\}=\frac{(\lambda p)^y e^{-\lambda}}{y!}e^{\lambda q}=\frac{(\lambda p)^y}{y!}e^{-\lambda p},\quad y=0,1,\cdots,$$

故可以确定一只昆虫所生幼虫数服从泊松分布 $\pi(\lambda p)$.

例 11 设随机点落在以原点为圆心、R 为半径的圆周上，并对极角是均匀分布的，求：

(1) 该点横坐标的概率密度；

(2) 求连接这点与点 $(-R,0)$ 的弦长的概率密度.

解 (1)设极角为随机变量 $\Theta\sim U(-\pi,\pi)$，圆周上任一点 A 的横坐标 X 取值为 x，$X=R\cos\Theta$，于是

$$f_\Theta(\theta)=\begin{cases}1/2\pi, & -\pi\leqslant\theta\leqslant\pi,\\ 0, & \text{其它}.\end{cases}$$

$x=g(\theta)$的反函数$\theta=h(x)=\arccos(x/R)$，当θ在$[-\pi,\pi]$上取值时，x在$[-R,R]$上取值，但$x=g(\theta)$不是单调函数，只能用分布函数法求$F_X(x)$. 由图 2.5 知

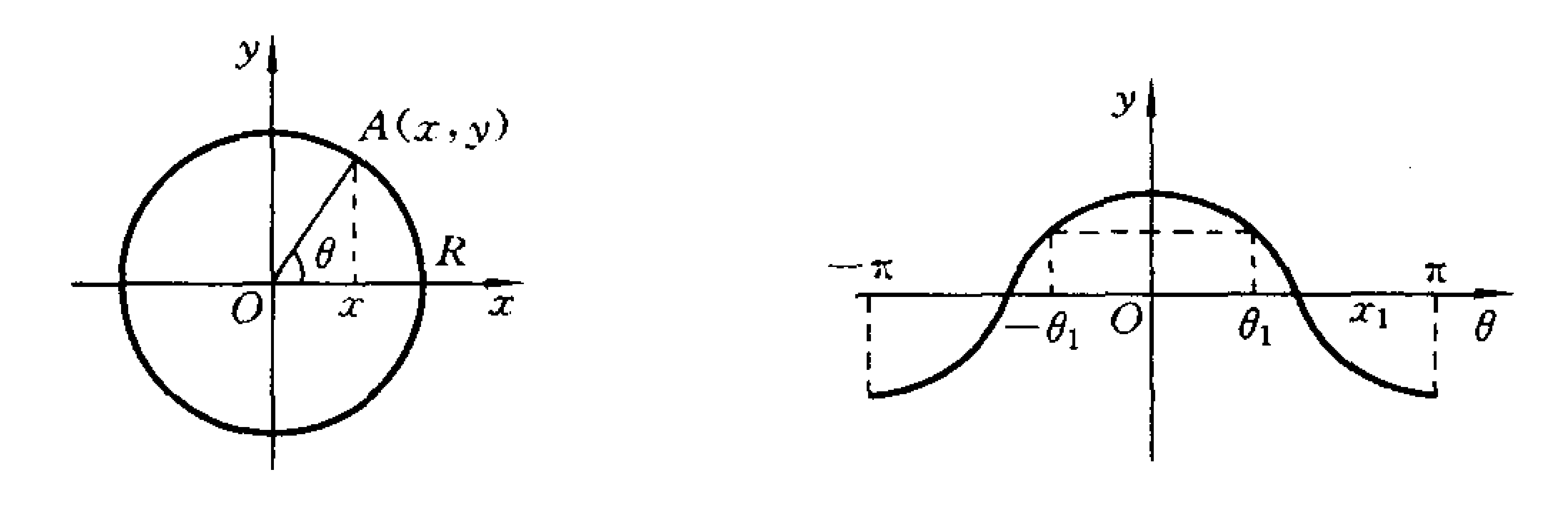


图 2.5

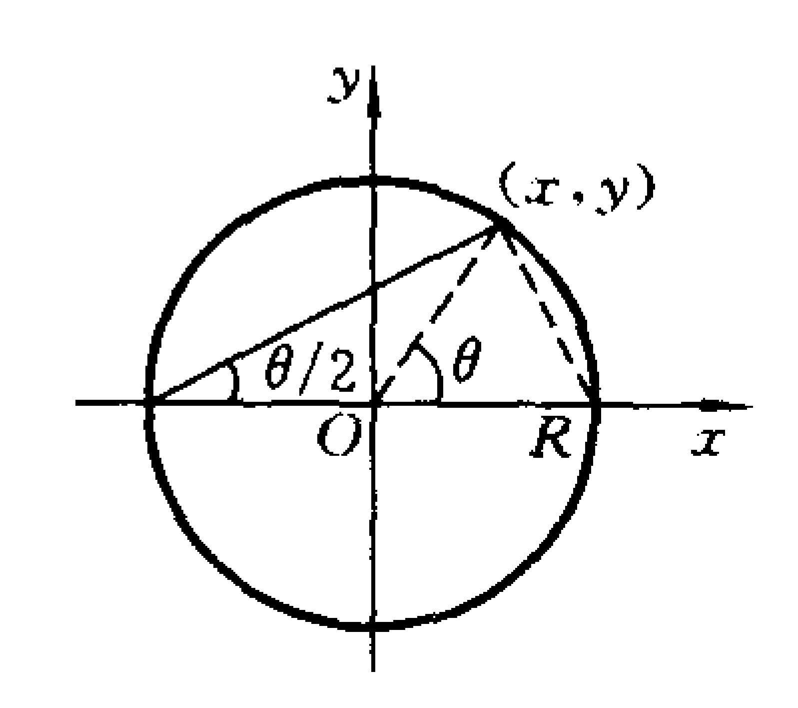
$$\begin{aligned}F_X(x)&=P\{X\leqslant x\}=P\{-\pi<\theta<-\theta_1\}+P\{\theta_1<\theta<\pi\}\\&=\int_{-\pi}^{-\theta_1}f(\theta)\mathrm{d}\theta+\int_{\theta_1}^{\pi}f(\theta)\mathrm{d}\theta\ \left(\text{取}\ \theta_1=\arccos\frac{x}{R}\right)\\&=\int_{-\pi}^{-\arccos(x/R)}\frac{1}{2\pi}\mathrm{d}\theta+\int_{\arccos(x/R)}^{\pi}\frac{1}{2\pi}\mathrm{d}\theta\\&=1-\frac{1}{\pi}\arccos\frac{x}{R},\quad -R<x<R,\end{aligned}$$

所以
$$f_X(x)=\begin{cases}1/(\pi\sqrt{R^2-x^2}), & -R<x<R,\\ 0, & \text{其它}.\end{cases}$$

(2) 设弦长为随机变量Z，取值$z\in[0,2R]$. $z=2R\cos\dfrac{\theta}{2}$，反函数$\theta=h(z)=2\arccos z(z/2R)$. 由图 2.6 知

$$\begin{aligned}F_Z&=P\{Z\leqslant z\}=P\{-\pi<\theta<-\theta_1\}+P\{\theta_1<\theta<\pi\}\\&=\int_{-\pi}^{-2\arccos(z/2R)}\frac{1}{2\pi}\mathrm{d}\theta+\int_{2\arccos(z/2R)}^{\pi}\frac{1}{2\pi}\mathrm{d}\theta\\&=1-\frac{2}{\pi}\arccos\frac{z}{2R},\quad 0\leqslant z\leqslant 2R,\end{aligned}$$

所以
$$f_Z(z)=\begin{cases}2/(\pi\sqrt{4R^2-z^2}), & 0\leqslant z\leqslant 2R,\\ 0, & \text{其它}.\end{cases}$$



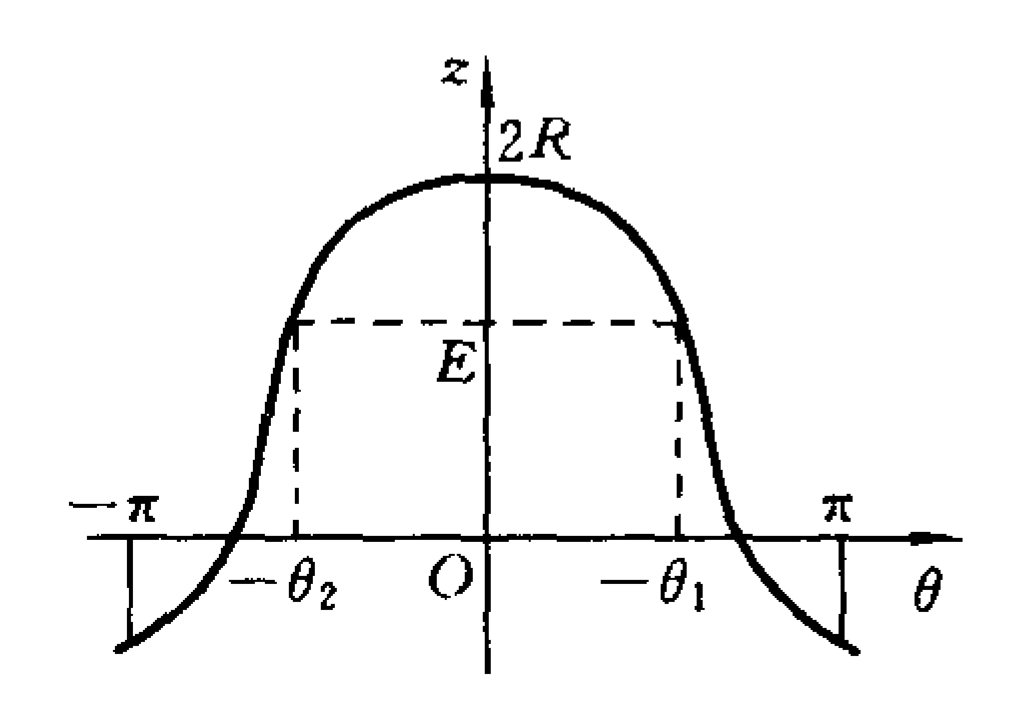


图 2.6

例 12 设 X 为连续型随机变量，分布函数为 $F(x)$，密度函数为 $f(x)$ 求下列随机变量的分布函数与概率密度：

(1) $Y_1=1/X$； (2) $Y_2=|X|$； (3) $Y_3=e^{-x}$.

解 (1) 因为 $F_{Y_1}(y_1)=P\{1/X<y_1\}$，所以

当 $y_1<0$ 时，由于 $\{1/X<y_1\}$ 等价于 $\{1/y_1<X<0\}$，于是

$$F_{Y_1}(y_1)=P\{1/y_1<X<0\}=F(0)-F(1/y_1).$$

当 $y_1=0$ 时，由于 $\{1/X<y_1\}$ 等价于 $\{X<0\}$，于是 $F_{Y_1}(y_1)=F(0)$.

当 $y_1>0$ 时，由于 $\{1/X>y_1\}$ 等价于 $\{X<0\}\cup\{X>1/y\}$，于是

$$F_{Y_1}(y_1)=F(0)+1-F(1/y_1).$$

又 $x=1/y_1$，$x'_y=-1/y_1^2<0$，由公式得

$$f_{Y_1}(y_1)=f\left(\frac{1}{y_1}\right)\cdot\frac{1}{y_1^2}=f\left(\frac{1}{y_1}\right)\Big/y_1^2\quad(y_1\neq0).$$

(2) 因为 $F_{Y_2}(y_2)=P\{|X|\leqslant y_2\}=P\{-y_2<X<y_2\}$，所以

当 $y_2>0$ 时，

$$F_{Y_2}(y_2)=F(y_2)-F(-y_2).$$

当 $y_2<0$ 时， $F_{Y_2}(y_2)=0$.

因此 $f_{Y_2}(y_2)=F'_{Y_2}(y_2)=\begin{cases}f(y_2)+f(-y_2), & y_2>0,\\ 0, & y_2<0.\end{cases}$

(3) 因为 $Y_3=e^{-X}$，所以 $X=-\ln y_3$. 当 $y_3>0$ 时，

$F_{Y_3}(y_3)=P\{e^{-X}<y_3\}=P\{X>-\ln y_3\}=1-F(-\ln y_3)$.

当 $y_3<0$ 时， $F_{Y_3}(y_3)=0$.

因此 $f_{Y_3}(y_3)=F'_{Y_3}(y_3)=\begin{cases} f(-\ln y_3)/y_3, & y_3>0, \\ 0, & y_3<0. \end{cases}$

例 13 设有一电路如图 2.7 所示. 电阻 R 是一随机变量，均匀分布在 900～1100 Ω 之间，电流 $i=0.01$ A，$r_0=1000$ Ω，求电压 $V=iR+ir_0$ 的概率密度.

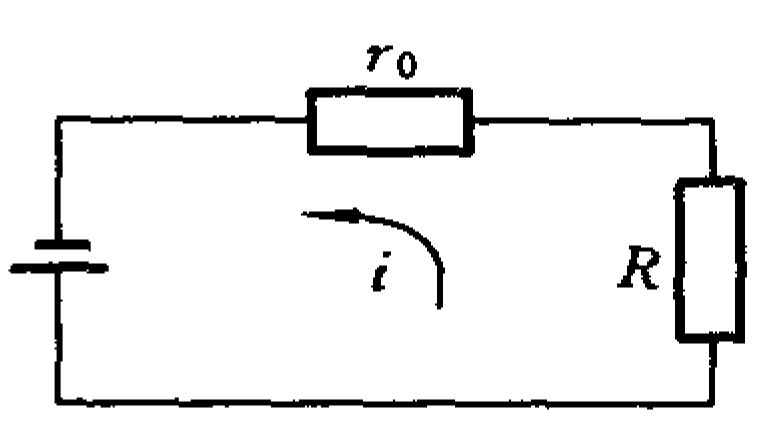


图 2.7

解 因为 $R\sim V(900,1100)$，$V=0.01R+10$，所以 V 也服从均匀分布，分布区间为

$$[0.01\times 900+10,0.01\times 1100+10]=[19,21],$$

所以 $$f_V(v)=\begin{cases} 1/2, & 19<V<21, \\ 0, & \text{其它}. \end{cases}$$

例 14 设 X 为连续型随机变量，若 (1) X 的概率密度为 $f_X(x)$，(2) $X\sim e(\lambda)$，求 $Y=X^3$ 的概率密度.

解 因 $Y=X^3$，而 $y=x^3$ 是单调增函数，可导，且反函数 $x=y^{1/3}$，$x'=\dfrac{1}{3}y^{-2/3}$，故

(1) $f_Y(y)=f_X(y^{1/3})\left(\dfrac{1}{3}y^{-2/3}\right)=\dfrac{1}{3}f(\sqrt[3]{y})\left(\dfrac{1}{\sqrt[3]{y^2}}\right)$， $y\neq 0$.

(2) 由 $f_X(x)$ 当 $x>0$ 时为 $\lambda e^{-\lambda x}$，$x\leqslant 0$ 时为零，得

$$f_Y(y)=\begin{cases} \dfrac{1}{3}\lambda e^{-\lambda\sqrt[3]{y}}/\sqrt[3]{y^2}, & y>0, \\ 0, & y\leqslant 0. \end{cases}$$

例 15 设某工程队完成某项工程所需时间 X（单位:d）近似服从 $N(100,5^2)$. 工程队上级规定：若工程在 100 d 内完工，可获奖金 10 万元；在 100～115 d 内完工，可获奖金 3 万元；超过 115 d 完工，罚款 5 万元. 求该工程队在完成此项工程时所获奖金的分布律.

解 $X\sim N(100,5^2)$，Y 是 X 的函数，可取值为 10，3，−5，故

$$P\{Y=-5\}=P\{115<X<+\infty\}=1-\Phi\left(\frac{115-100}{5}\right)$$

$$=1-\Phi(3)=0.0013,$$

$$P\{Y=3\}=P\{100<X\leqslant 115\}$$

$$=\Phi\left(\frac{115-100}{5}\right)-\Phi\left(\frac{100-100}{5}\right)$$

$$=\Phi(3)-\Phi(0)=0.4987,$$

$$P\{Y=10\}=P\{X\leqslant 100\}=\Phi\left(\frac{100-100}{5}\right)$$

$$=\Phi(0)=0.5000.$$

所以，所获奖金Y的分布律为

Y	-5	3	10
p_k	0.0013	0.4987	0.5000

故从本例得知，连续型随机变量的函数也可以是离散型的.

例16 设电流I是一个随机变量，均匀分布在9～11 A之间.若此电流通过2 Ω的电阻，在电阻上消耗功率$W=2I^2$，求W的概率密度.

解 反函数$I=\pm\sqrt{W/2}$，$162\leqslant W\leqslant 242$. 当$W>0$时，

$$F_W(w)=P\{W\leqslant w\}=\int_{-\sqrt{w/2}}^{\sqrt{w/2}}f(I)\mathrm{d}I,$$

当$W<0$时，$F_W(w)=0$.

由于
$$f_I(i)=\begin{cases}1/2, & i\in(9,11),\\ 0, & \text{其它},\end{cases}$$

所以
$$f_W(w)=\begin{cases}\dfrac{1}{8}\sqrt{2/w}, & 162<W<242,\\ 0, & \text{其它}.\end{cases}$$

硕士研究生入学试题分析

一、本章考试要求

1. 理解随机变量及其概率分布的概念，理解分布函数($F(x)$

$=P\{X\leqslant x\}$)的概念及性质，会计算与随机变量有关的事件的概率.

2. 理解离散型随机变量及其概率分布的概念，掌握0-1分布、二项分布、超几何分布、泊松分布及其应用.

3. 理解连续型随机变量及其概率密度的概念，掌握概率密度

$$f(x)=\begin{cases}\lambda e^{-\lambda x}, & x>0,\\ 0, & x\leqslant 0\end{cases}$$

与分布函数之间的关系，掌握正态分布、均匀分布、指数分布及其应用.

4. 会求简单随机变量函数的概率分布.

二、本章重点内容

求随机变量相关事件的概率，随机变量的分布函数或概率密度，求随机变量函数的分布. 在硕士研究生入学试题中很可能与随机变量的数字特征构成综合题.

（一）离散型随机变量的概率分布

1. 从数1,2,3,4中任取一个数，记为X，再从$1,2,\cdots,X$中任取一个数，记为Y，则$P\{Y=2\}=$________.　　（2005年三、四）

解　因为

$$P\{X=i\}=1/4\ (i=1,2,3,4),$$

而　$$P\{Y=2|X=1\}=0,\quad P\{Y=2|X=2\}=1/2,$$

$$P\{Y=2|X=3\}=1/3,\quad P\{Y=2|X=4\}=1/4,$$

故　$$P\{Y=2\}=1/4\times(0+1/2+1/3+1/4)=13/48.$$

2. 设随机变量X服从正态分布$N(0,1)$，对给定的$\alpha\ (0<\alpha<1)$，数u_α满足$P\{X>u_\alpha\}=\alpha$. 若$P\{|X|<x\}=\alpha$，则x等于(　　).

(A) $u_{\alpha/2}$;　　(B) $u_{1-\alpha/2}$;　　(C) $u_{(1-\alpha)/2}$;　　(D) $u_{1-\alpha}$.

（2004年一、三、四）

解　选(C). 由$P\{|X|<x\}=\alpha$知，$P\{|X|\geqslant x\}=1-\alpha$. 又由分布$N(0,1)$的对称性知，$x=u_{(1-\alpha)/2}$.

3. 设随机变量$X\sim B(2,p)$，$Y\sim B(3,p)$，若$P\{X\geqslant 1\}=5/9$，

则 $P\{Y\geqslant 1\}=$ ________.　　　　　　　　　　　　　　(1997 年四)

解　本题的关键是先求出 p. 解方程

$$P\{X\geqslant 1\}=C_2^1p(1-p)+C_2^2p^2(1-p)^0=2p-p^2=5/9,$$

得 $p=1/3$(负值舍去),于是

$$P\{Y\geqslant 1\}=1-C_3^0\times(1/3)^0\times(1-1/3)^3=19/27.$$

4. 假设一厂家生产的每台仪器,以概率0.7可以直接出厂,以概率0.3需进一步调试. 经调试后以概率0.8可以出厂,以概率0.2定为不合格品不能出厂. 现该厂生产了 n $(n\geqslant 2)$ 台仪器(假设各台仪器的生产过程相互独立). 求:

(1) 全部能出厂的概率 α;

(2) 其中恰好有两台不能出厂的概率 β;

(3) 其中至少有两台不能出厂的概率 θ.　　　　　　　　(1995 年四)

解　记 $A=\{$仪器需进一步调试$\}$,记 $B=\{$仪器能出厂$\}$,则 $\overline{A}=\{$仪器能直接出厂$\}$,$AB=\{$仪器经调试后能出厂$\}$.

由条件知,$B=\overline{A}+AB$,且

$$P(A)=0.3,\quad P(B|A)=0.8,$$

$$P(AB)=P(A)P(B|A)=0.3\times 0.8=0.24,$$

$$P(B)=P(\overline{A})+P(AB)=0.7+0.24=0.94.$$

设 X 为所生产的 n 台仪器中能出厂的台数,则 $X\sim B(n,p)=B(n,0.94)$,所以

$$\alpha=P\{X=n\}=0.94^n,$$

$$\beta=P\{X=n-2\}=C_n^2\times 0.94^{n-2}\times 0.06^2,$$

$$\theta=P\{X\leqslant n-2\}=1-P\{x=n-1\}-P\{x=n\}$$
$$=1-n\times 0.94^{n-1}\times 0.06-0.94^n.$$

5. 设 $F_1(x)$ 与 $F_2(x)$ 分别为随机变量 X_1 与 X_2 的分布函数,为使 $F(x)=aF_1(x)-bF_2(x)$ 是某一随机变量的分布函数,在下列给定的各组数值中应取(　　).

(A) $a=3/5,b=-2/5$;　　　　　(B) $a=2/3,b=2/3$;

(C) $a=-1/2, b=3/2$;　　　　(D) $a=1/2, b=-3/2$.

（1998 年四）

解　选(A).

$$\lim_{x\to+\infty} F_1(x)=1,\quad \lim_{x\to+\infty} F_2(x)=1,$$

$$\lim_{n\to+\infty} aF_1(x)-bF_2(x)=a-b=1,$$

故(A)满足此要求.

（二）连续型随机变量的概率分布

1. 设随机变量 X 的概率密度为

$$f(x)=\begin{cases}1/(3\sqrt[3]{x^2}), & x\in[1,8],\\ 0, & 其它,\end{cases}$$

$F(x)$是 X 的分布函数,求随机变量$Y=F(X)$的分布函数.

（2003 年三、四）

解　当 $x<1$ 时,$F(x)=0$;当 $x>8$ 时,$F(x)=1$;而当 $x\in[1,8]$时,$F(x)=\int_1^x 1/(3\sqrt[3]{x^2})\mathrm{d}x=\sqrt[3]{x}-1$.

设$Y=F(X)$的分布函数为$G(y)$,则:当$y\leqslant 0$ 时,$G(y)=0$;当$y\geqslant 1$ 时,$G(y)=1$;当 $y\in(0,1)$时,有

$$\begin{aligned}G(y)&=P\{Y\leqslant y\}=P\{F(X)\leqslant y\}\\&=P\{\sqrt[3]{X}-1\leqslant y\}=P\{X\leqslant(y+1)^3\}\\&=F[(y+1)^3]=\sqrt{(y+1)^3}-1=y,\end{aligned}$$

所以,$Y=F(X)$的分布函数为

$$G(y)=\begin{cases}0, & y\leqslant 0,\\ y, & 0<y<1,\\ 1, & y\geqslant 1.\end{cases}$$

2. 设 X_1 和 X_2 是任意两个相互独立的连续型随机变量,它们的概率密度分别为 $f_1(x)$和 $f_2(x)$,分布函数分别为 $F_1(x)$和 $F_2(x)$,则(　　).

(A) $f_1(x)+f_2(x)$必为某一随机变量的概率密度;

(B) $f_1(x)f_2(x)$必为某一随机变量的概率密度；

(C) $F_1(x)+F_2(x)$必为某一随机变量的分布函数；

(D) $F_1(x)F_2(x)$必为某一随机变量的分布函数.（2002 年一）

解 选(D).

(1) $\int_{-\infty}^{+\infty}[f_1(x)+f_2(x)]\mathrm{d}x=2$，所以(A)不成立.

(2) 设$X_1\sim U(0,2)$，$X_2\sim U(0,4)$，则

$$\int_{-\infty}^{+\infty}f_1(x)f_2(x)\mathrm{d}x=\int_0^2\frac{1}{2}\times\frac{1}{4}\mathrm{d}x=\frac{1}{4},$$

所以(B)不成立.

(3) $F(+\infty)=F_1(+\infty)+F_2(+\infty)=2$，所以(C)不成立.

(4) 可以验证，$0\leqslant F_1(x)F_2(x)\leqslant 1$，单调不减，右连续，所以(D)成立.

3. 设随机变量X与Y均服从正态分布，$X\sim N(\mu,4^2)$，$Y\sim N(\mu,5^2)$，记$p_1=P\{X\leqslant\mu-4\}$，$p_2=P\{Y\geqslant\mu+5\}$，则( ).

(A) 对任何实数μ，都有$p_1=p_2$；

(B) 对任何实数μ，都有$p_1<p_2$；

(C) 只对μ的个别值，才有$p_1=p_2$；

(D) 对任何实数μ，都有$p_1>p_2$.（1993 年五）

解 选(A).化为标准正态分布讨论.

$$p_1=P\{X\leqslant\mu-4\}=P\left\{\frac{X-\mu}{4}\leqslant\frac{\mu-4-\mu}{4}\right\}=P\left\{\frac{X-\mu}{4}\leqslant-1\right\},$$

$$p_2=P\{Y\geqslant\mu+5\}=P\left\{\frac{X-\mu}{5}\geqslant\frac{\mu+5-\mu}{5}\right\}=P\left\{\frac{Y-\mu}{5}\geqslant 1\right\},$$

由标准正态分布的对称性知，$p_1=p_2$.

4. 设随机变量X服从正态分布$N(\mu,\sigma^2)$，则随σ的增大，概率$P\{|X-\mu|<\sigma\}$( ).

(A) 单调增大； (B) 单调减小；

(C) 保持不变； (D) 增减不定.（1995 年四）

解 选(C).因为

$$P\{|X-\mu|<\sigma\}=P\left\{\frac{|X-\mu|}{\sigma}<\frac{\sigma}{\sigma}\right\}=P\left\{\frac{|X-\mu|}{\sigma}<1\right\},$$

所以不因σ的改变而改变.

5. 设随机变量X的密度函数为$\varphi(x)$，且$\varphi(x)=\varphi(-x)$，$F(x)$是X的分布函数，则对任意实数a，有(　　).

(A) $F(-a)=1-\int_0^a\varphi(x)\mathrm{d}x$； (B) $F(-a)=\frac{1}{2}-\int_0^a\varphi(x)\mathrm{d}x$；

(C) $F(-a)=F(a)$； (D) $F(-a)=2F(a)-1$.

(1993年四)

解 选(B). 由$\varphi(-x)=\varphi(x)$知，密度函数曲线关于y轴对称，故

$$F(-a)=\frac{1}{2}-\int_0^a\varphi(x)\mathrm{d}x.$$

6. 设随机变量X服从指数分布，则随机变量$Y=\min(X,2)$的分布函数(　　).

(A) 是连续函数； (B) 至少有两个间断点；

(C) 是阶跃函数； (D) 恰好有一个间断点.

(1999年四)

解 选(D). X是连续函数，但在$Y=2$处，Y的分布函数间断，于是有$F(y)=\begin{cases}1-\mathrm{e}^{-\lambda y}, & y<2,\\ 1, & y\geqslant 2.\end{cases}$

7. 设随机变量X的概率密度为

$$f(x)=\begin{cases}1/3, & 若\ x\in[0,1],\\ 2/9, & 若\ x\in[3,6],\\ 0, & 其它.\end{cases}$$

若k使得$P\{X\geqslant k\}=2/3$，则k的取值范围是______. (2000年三)

解 [1,3]. 因$P\{X\geqslant k\}=2/3$，故$P\{X<k\}=1/3$. 由于

$$P\{X\leqslant 1\}=\int_0^1\frac{1}{3}\mathrm{d}x=\frac{1}{3},$$

故k取值范围是[1,3].

8. 设随机变量 X 服从(0,2)上的均匀分布，则随机变量 $Y=X^2$ 在(0,4)内的概率分布密度 $f_Y(y)=$________. (1993 年一)

解 $$f_X(x)=\begin{cases}1/2, & x\in(0,2),\\ 0, & 其它,\end{cases}$$

而 $y=x^2, x=\sqrt{y}, x'=\dfrac{1}{2\sqrt{y}}$单调. 由定理有

$$f_Y(y)=\begin{cases}\dfrac{1}{2}\times\dfrac{1}{2\sqrt{y}}=\dfrac{1}{4\sqrt{y}}, & (0,4),\\ 0, & 其它.\end{cases}$$

9. 已知随机变量 X 的概率密度为

$$f(x)=\frac{1}{2}e^{-|x|},\quad -\infty<x<+\infty,$$

则 X 的分布函数 $F(x)=$________. (1990 年一)

解 $$F(x)=\begin{cases}\displaystyle\int_{-\infty}^{x}\frac{1}{2}e^{t}dt=\frac{1}{2}e^{x}, & x<0,\\ \displaystyle\int_{-\infty}^{0}\frac{1}{2}e^{t}dt+\int_{0}^{x}\frac{1}{2}e^{-t}dt=1-\frac{1}{2}e^{-x}, & x>0.\end{cases}$$

10. 设随机变量X 服从均值为10，均方差为0.02 的正态分布. 已知

$$\Phi(x)=\int_{-\infty}^{x}\frac{1}{\sqrt{2\pi}}e^{-u^2/2}du,\quad \Phi(2.5)=0.9938,$$

则 X 落在(9.95,10.05)内的概率为________. (1988 年一)

解 $P\{9.95<x<10.05\}$

$$=P\left\{\frac{9.95-10}{0.02}<\frac{X-10}{0.02}<\frac{10.05-10}{0.02}\right\}$$

$$=2\Phi(2.5)-1=0.9876.$$

11. 设随机变量 X 的概率密度为

$$f_X(x)=\begin{cases}e^{-x}, & x\geqslant0,\\ 0, & x<0,\end{cases}$$

求随机变量 $Y=e^X$ 的概率密度 $f_Y(y)$. (1995 年一)

解 $F_Y(y)=P\{Y<y\}=P\{e^X<y\}$

$$=\begin{cases}P\{X<\ln y\}, & y\geqslant 1,\\ 0, & y<1.\end{cases}$$

当 $y\geqslant 1$ 时，$\quad F_Y(y)=P\{X<\ln y\}=\int_0^{\ln y}\mathrm{e}^{-x}\mathrm{d}x,$

因此 $\quad f_Y(y)=\begin{cases}0, & y<1,\\ F'_Y(y)=1/y^2, & y\geqslant 1.\end{cases}$

12. 假设随机变量 X 服从参数为 2 的指数分布. 证明 $Y=1-\mathrm{e}^{-2X}$在区间(0,1)内服从均匀分布. (1995 年五)

解 $\quad F_x=\begin{cases}1-\mathrm{e}^{-2x}, & x>0,\\ 0, & x\leqslant 0.\end{cases}$

$y=1-\mathrm{e}^{-2x}$是单调增函数，$x=\frac{1}{2}\ln(1-y)$，则

$$G(y)=P\{Y\leqslant y\}=p\{1-\mathrm{e}^{2X}\leqslant y\}$$

$$=\begin{cases}0,\\ P\left\{X\leqslant -\frac{1}{2}\ln(1-y)\right\},\\ 1,\end{cases}$$

故 $\quad G(y)=\begin{cases}0, & y\leqslant 0,\\ y, & 0<y<1,\\ 1, & y\geqslant 1.\end{cases}$

13. 设随机变量 X 的概率密度 $f(x)=\begin{cases}2x, & (0,1),\\ 0, & 其它,\end{cases}$ 以 Y 表示对 X 的三次独立重复观察中，事件$\{X\leqslant 1/2\}$出现的次数，则 $P\{Y=2\}=$________. (1994 年四)

解 $P\left\{X\leqslant\frac{1}{2}\right\}=\int_0^{1/2}2x\mathrm{d}x=x^2\Big|_0^{1/2}=\frac{1}{4}$，故

$$P(Y=2)=C_3^2\times\left(\frac{1}{4}\right)^2\times\frac{3}{4}=\frac{9}{64}.$$

14. 设随机变量 X 的概率密度 $f(x)=\begin{cases}2x, & (0,1),\\ 0, & 其它,\end{cases}$ 现对 X 进行n 次独立重复观察，以V_n 表示观察值不大于0.1 的次数，试求随机变量 V_n 的概率分布. (1994 年五)

解 $p=P\{X\leqslant 0.1\}=\int_{-\infty}^{0.1}f(x)\mathrm{d}x=2\int_{0}^{0.1}x\mathrm{d}x=0.01.$

记 $X\leqslant 0.1$ 为成功，则 $V_n\sim B(n,0.01)$，有

$$P\{V_n=m\}=C_n^m\times 0.01^m\times 0.99^{n-m}\quad (m=0,1,\cdots,n).$$

15. 假设一电路装有三个同种电气元件，其工作状态相互独立，且无故障工作时间都服从参数为 λ $(\lambda>0)$ 的指数分布. 当三个元件都无故障时，电路正常工作，否则整个电路不能正常工作，试求电路正常工作时间 T 的概率分布. （1996 年五）

解 以 X_i $(i=1,2,3)$ 表示第 i 个电气元件无故障工作的时间，则 X_1,X_2,X_3 相互独立且同分布. 分布函数为

$$F(x)=\begin{cases}1-\mathrm{e}^{-\lambda x}, & x>0,\\ 0, & x\leqslant 0.\end{cases}$$

设 $G(t)$ 是 T 的分布函数. 当 $t\leqslant 0$ 时，$G(t)=0$；当 $t>0$ 时，有

$$\begin{aligned}G(t)&=P\{T\leqslant t\}=1-P\{T>t\}\\&=1-P\{X_1>t,X_2>t,X_3>t\}\\&=1-P\{X_1>t\}P\{X_2>t\}P\{X_3>t\}\\&=1-[1-F(t)]^3=1-\mathrm{e}^{-3\lambda t},\end{aligned}$$

故

$$G(t)=\begin{cases}0, & t\leqslant 0,\\ 1-\mathrm{e}^{-3\lambda t}, & t>0,\end{cases}$$

即 $T\sim \mathrm{e}(3\lambda)$.

16. 设随机变量 X 的绝对值不大于 1，$P\{X=-1\}=1/8$，$P\{X=1\}=1/4$. 在事件 $\{-1<X<1\}$ 出现的条件下，X 在 $(-1,1)$ 内的任一子区间上取值的条件概率与该子区间长度成正比，试求 X 的分布函数 $F(x)=P\{X\leqslant x\}$. （1997 年三）

解 显然，当 $x<-1$ 时，$F(x)=0$；而当 $x\geqslant 1$ 时，$F(x)=1$. 又 $|X|\leqslant 1$，故 $P\{-1<X<1\}=1-1/8-1/4=5/8$，则由均匀分布的定义知

$$P\{-1<X\leqslant x\,|\,-1<X<1\}=(X+1)/2,$$

$$P\{-1<X\leqslant x\}$$

$$=P\{-1<X\leqslant x,-1<X<1\}$$
$$=P\{-1<X<1\}P\{-1<X\leqslant x\mid -1<X<1\}$$
$$=\frac{5}{8}\times\frac{x+1}{2}=\frac{5(x+1)}{16}.$$

即当 $-1\leqslant X<1$ 时，

$$F(x)=P\{X\leqslant -1\}+P\{-1<X\leqslant x\}$$
$$=1/8+5(x+1)/16,$$

故
$$F(x)=\begin{cases}0, & x<-1,\\ 1/8+5(x+1)/16, & -1\leqslant x<1,\\ 1, & x\geqslant 1.\end{cases}$$

17. 设随机变量 X 的概率密度函数为

$$f_X(x)=1/[\pi(1+x^2)],\quad -\infty<x<+\infty,$$

求随机变量 $Y=1-\sqrt[3]{X}$ 的概率密度函数. (1988 年一)

解 用分布函数法求. 由

$$P\{Y\leqslant y\}=P\{1-\sqrt[3]{X}\leqslant y\}=P\{X\geqslant(1-y)^3\},$$

得
$$F_Y(y)=P\{X\geqslant(1-y)^3\}=1-P\{X\leqslant(1-y)^3\}$$
$$=1-\int_{-\infty}^{(1-y)^3}\frac{1}{\pi(1+x^2)}\mathrm{d}x,$$
$$f_Y(y)=F_Y'(y)=\frac{3(1-y)^2}{\pi[1+(1-y)^6]},\ -\infty<y<+\infty.$$

18. 设随机变量 X 在 $[1,6]$ 上服从均匀分布，则方程 $x^2+Xx+1=0$ 有实根的概率是________. (1989 年一)

解 因为 $\Delta=X^2-4\geqslant 0$ 成立，$X\geqslant 2$，所以

$$p=(6-2)/(6-1)=0.8.$$

19. 若 $X\sim N(2,\sigma^2)$，且 $P\{2<X<4\}=0.3$，求 $P\{X<0\}$.

(1991 年一)

解 $P\{X<0\}=P\left\{\frac{X-2}{\sigma}<\frac{-2}{\sigma}\right\}=1-\Phi\left(\frac{2}{\sigma}\right)$，而

$$P\{2<X<4\}=P\left\{\frac{2-2}{\sigma}<\frac{X-2}{\sigma}<\frac{4-2}{\sigma}\right\}=\Phi\left(\frac{2}{\sigma}\right)-\Phi(0),$$

即 $$\Phi\left(\frac{2}{\sigma}\right)=0.3+\Phi(0)=0.3+0.5=0.8.$$

于是 $$P\{X<0\}=1-0.8=0.2.$$

20. 设随机变量 X 服从正态分布 $N(\mu,\sigma^2)$ $(\sigma>0)$，且二次方程 $y^2+4y+X=0$ 无实根的概率为 1/2，则 $\mu=$ ________.

(2002 年一)

解 若二次方程无实根，则必有判别式$\Delta=4^2-4X<0$，即$X>4$. 所以，$P\{X>4\}=1/2$，于是，$P\{X<4\}=1/2$，知 $x=4$ 为正态分布 $N(\mu,\sigma^2)$的对称点，从而知 $\mu=4$.

21. 设X_1 和X_2 是任意两个相互独立的连续型随机变量，它们的概率密度分别为 $f_1(x)$和 $f_2(x)$，分布函数分别为 $F_1(x)$和 $F_2(x)$，则(　　).

(A) $f_1(x)+f_2(x)$必为某一随机变量的概率密度；

(B) $F_1(x)F_2(x)$必为某一随机变量的分布函数；

(C) $F_1(x)+F_2(x)$必为某一随机变量的分布函数；

(D) $f_1(x)f_2(x)$必为某一随机变量的概率密度. (2002 年四)

解 选(B).

因为$\int_{-\infty}^{+\infty}[f_1(x)+f_2(x)]\mathrm{d}x=2$，所以(A)不成立.

因为$\lim\limits_{x\to 0}[F_1(x)+F_2(x)]=2$，所以(C)不成立.

因为若 $X_1\sim U(0,1)$，$X_2\sim U(1,2)$，则 $f_1(x)$，$f_2(x)$不能成为随机变量的概率密度，所以(D)不成立.

而 $F_1(x)$，$F_2(x)$经过验证，满足不减性、有界性和右连续性，所以为某一随机变量的分布函数.

第三章　多维随机变量及其分布

第一节　二维随机变量及其概率分布

主要内容

1. 二维随机变量及其分布函数

设 $X_1,X_2,\cdots,X_n$ 是定义在同一样本空间 Ω 上的随机变量，则向量 $(X_1,X_2,\cdots,X_n)$ 称为 n 维随机变量或 n 维随机向量. 当 $n=2$ 时，称为二维随机变量，记为 (X,Y) 或 (ξ,η).

对于任意实数 x,y，二元函数

$$F(x,y)=P\{X\leqslant x,Y\leqslant y\}$$

称为二维随机变量 (X,Y) 的联合分布函数.

2. 联合分布函数 $F(x,y)$ 的性质

分布函数 $F(x,y)$ 在 (x,y) 处的值就是随机点 (X,Y) 落在图 3.1 所示以点 (x,y) 为顶点的位于其左下方的无穷矩形域内的概率.

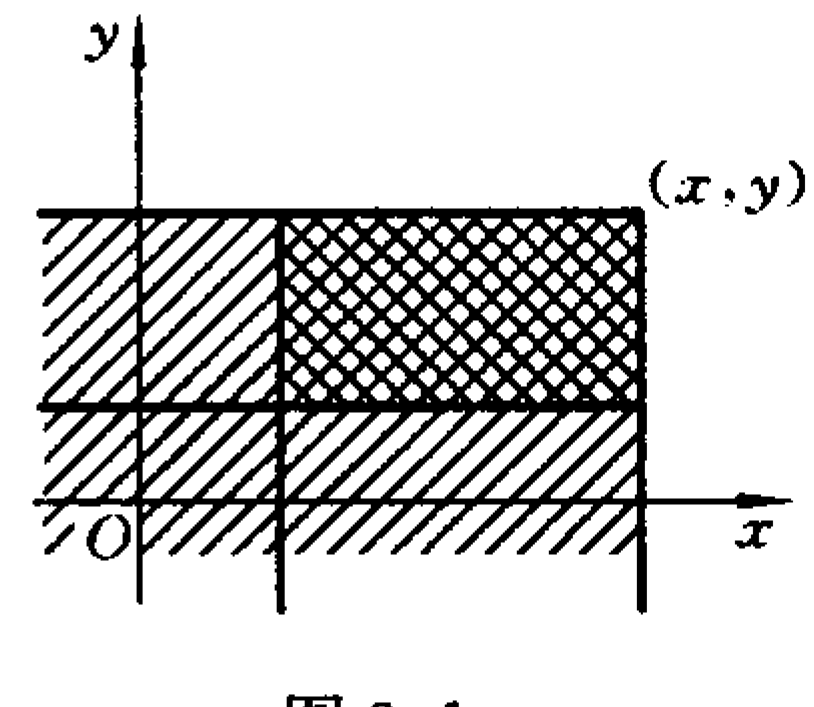


图 3.1

分布函数具有以下性质：

(1) $F(x,y)$ 是变量 x 和 y 的不减函数.

(2) $0\leqslant F(x,y)\leqslant 1$，且

$$\lim_{\substack{x\to-\infty,\\ y\text{固定}}} F(x,y)=0,\quad \lim_{\substack{y\to-\infty,\\ x\text{固定}}} F(x,y)=0,$$

$$\lim_{x\to-\infty,y\to-\infty} F(x,y)=0,\quad \lim_{x\to+\infty,y\to+\infty} F(x,y)=1.$$

(3) $F(x,y)$关于x右连续，关于y也右连续，即

$$F(x,y)=F(x+0,y),\quad F(x,y)=F(x,y+0).$$

(4) 对任意(x_1,y_1)，(x_2,y_2)，若$x_1<x_2$，$y_1<y_2$，则

$$F(x_2,y_2)-F(x_2,y_1)-F(x_1,y_2)+F(x_1,y_1)>0.$$

相当于随机点(X,Y)落在图3.1中画有交叉线的阴影区域内的概率（其右上角顶点为(x_2,y_2)，左下角顶点为(x_1,y_1)）.

3. 二维离散型随机变量及其概率分布

如果二维随机变量(X,Y)的所有可取值为有限对或可列无限多对，则称(X,Y)是离散型随机变量.

若(X,Y)的所有可能取值为(x_i,y_j)，$i,j=1,2,\cdots$，则称$P\{X=x_i,Y=y_j\}=p_{ij}$为二维离散型随机变量(X,Y)的概率分布或X和Y的联合分布律.

联合分布$P\{X=x_i,Y=y_j\}=p_{ij}$有以下性质：

$$p_{ij}\geqslant 0,\quad \sum_{i=1}^{\infty}\sum_{j=1}^{\infty}p_{ij}=1.$$

离散型随机变量(X,Y)的联合分布具有形式

$$F(x,y)=\sum_{x_i\leqslant x,\ y_j\leqslant y}\sum p_{ij}.$$

4. 二维连续型随机变量及其概率密度

设对二维随机变量(X,Y)的分布函数$F(x,y)$，存在非负函数$f(x,y)$，使得对于任何的x,y，有

$$F(x,y)=\int_{-\infty}^{y}\int_{-\infty}^{x}f(u,v)\mathrm{d}u\mathrm{d}v,$$

则(X,Y)为二维连续型随机变量. 函数$f(x,y)$称为(X,Y)的联合概率密度函数.

5. 概率密度函数$f(x,y)$的性质

(1) $f(x,y)\geqslant 0$，即$f(x,y)$为非负函数；

(2) $\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}f(x,y)\mathrm{d}x\mathrm{d}y=F(+\infty,+\infty)=1$；

(3) 在$f(x,y)$的连续点(x,y)，有$\dfrac{\partial^2 F(x,y)}{\partial x\partial y}=f(x,y)$；

(4) 对 xOy 平面上的任意区域 G，点 (X,Y) 落在 G 内的概率是

$$P\{(X,Y)\in G\}=\iint\limits_{G} f(x,y)\mathrm{d}x\mathrm{d}y.$$

疑难解析

1. 事件 $\{X\leqslant x,Y\leqslant y\}$ 表示事件 $\{X\leqslant x\}$ 与 $\{Y\leqslant y\}$ 的积事件，为什么 $P\{X\leqslant x,Y\leqslant y\}$ 不一定等于 $P\{X\leqslant x\}P\{Y\leqslant y\}$?

答 与仅当事件 A,B 相互独立时才有 $P(AB)=P(A)P(B)$ 一样. 这里，依乘法原理，有

$$P\{X\leqslant x,Y\leqslant y\}=P\{X\leqslant x\}P\{Y\leqslant y\mid X\leqslant x\},$$

那么，当 $P\{X\leqslant x\}$ 和 $P\{Y\leqslant y\}$ 相互独立时，有 $P\{Y\leqslant y\mid X\leqslant x\}=P\{Y\leqslant y\}$，因而下式成立：

$$P\{X\leqslant x,Y\leqslant y\}=P\{X\leqslant x\}P\{Y\leqslant y\}.$$

2. 事件 $\{X\leqslant a,Y\leqslant b\}$ 与事件 $\{X>a,Y>b\}$ 是否为对立事件? 为什么?

答 事件 $\{X\leqslant a,Y\leqslant b\}$ 与事件 $\{X>a,Y>b\}$ 不是对立事件，由图 3.2 知，事件 $\{X\leqslant a,Y\leqslant b\}$ 发生，即随机点 (X,Y) 落在图左下部阴影区域内；事件 $\{X>a,Y>b\}$ 发生，即随机点 (X,Y) 落在图右上部阴影区域内. 它们的和事件不遮盖全平面区域，所以不是对立事件.

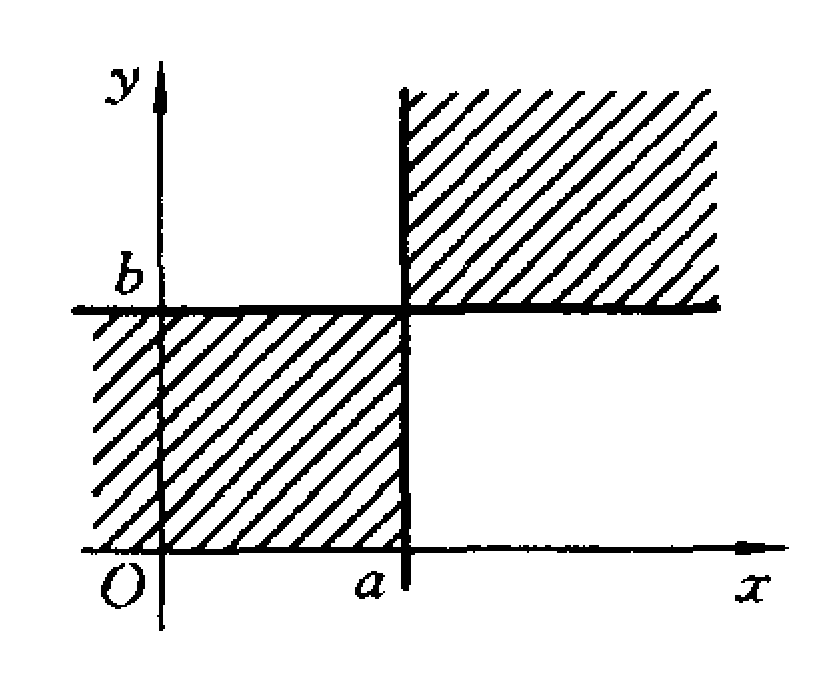


图 3.2

3. 计算概率 $P\{(X,Y)\in G\}$ 时应注意些什么?

答 当 (X,Y) 是离散型随机变量时，

$$P\{(X,Y)\in G\}=\sum_{(x_i,y_j)\in G}\sum p_{ij}.$$

注意，必须找出 G 内所有使 $p_{ij}\neq 0$ 的点 (x_i,y_j)，不能遗漏.

当 (X,Y) 是连续型随机变量时，

$$P\{(X,Y)\in G\}=\iint\limits_{G} f(x,y)\mathrm{d}x\mathrm{d}y.$$

注意，必须正确确定二重积分的积分限. 特别是在区域G要分块计算时，不要有遗漏或重复求积的现象发生.

在求(X,Y)的分布函数$F(x,y)$时，对$\mathbf{R}^2$平面上的各个区域的概率都要正确求出，并按顺序累积(可运用矩形区域的和逆等运算，求得$F(x,y)$).

方法、技巧与典型例题分析

一、二维离散型随机变量(X,Y)的联合分布的求法

在理解二维随机变量和二维随机变量的分布函数等概念的基础上，确定实际问题中的(X,Y)的所有可能取值. 通常可以由定义与古典型概率方式求出$P\{X=x_i,Y=y_j\}$；但有时要借助于事件的关系与运算性质，如和、差、积、独立性与全概率公式等来求. 将全部$P\{X=x_i,Y=y_j\}$列出，即得(X,Y)的分布律.

二、二维离散型随机变量的分布函数的求法

求二维离散型随机变量的分布函数时，不仅要正确计算p_{ij}，更要注意验证是否满足分布函数的性质.

例1 在元旦茶话会上，每人发给一袋水果，内装3只橘子，2只苹果，3只香蕉. 今从袋中随机抽出4只，以X记橘子数，Y记苹果数，求(X,Y)的分布律.

解 X可取值为0,1,2,3，Y可取值0,1,2.

$$P\{X=0,Y=0\}=P\{\varnothing\}=0,$$
$$P\{X=0,Y=1\}=C_3^0C_2^1C_3^3/C_8^4=2/70,$$
$$P\{X=0,Y=2\}=C_3^0C_2^2C_3^2/C_8^4=3/70,$$
$$P\{X=1,Y=0\}=C_3^1C_2^0C_3^3/C_8^4=3/70,$$
$$P\{X=1,Y=1\}=C_3^1C_2^1C_3^2/C_8^4=18/70,$$
$$P\{X=1,Y=2\}=C_3^1C_2^2C_3^1/C_8^4=9/70,$$

$$P\{X=2,Y=0\}=C_3^2C_2^0C_3^2/C_8^4=9/70,$$
$$P\{X=2,Y=1\}=C_3^2C_2^1C_3^1/C_8^4=18/70,$$
$$P\{X=2,Y=2\}=C_3^2C_2^2C_3^0/C_8^4=3/70,$$
$$P\{X=3,Y=0\}=C_3^3C_2^0C_3^1/C_8^4=3/70,$$
$$P\{X=3,Y=1\}=C_3^3C_2^1C_3^0/C_8^4=2/70,$$
$$P\{X=3,Y=2\}=P(\varnothing)=0.$$

所以，(X,Y)的联合分布律如下：

Y \ X	0	1	2	3
0	0	3/70	9/70	3/70
1	2/70	18/70	18/70	2/70
2	3/70	9/70	3/70	0

例 2 将一枚均匀硬币掷三次，以 X 记正面出现的次数，以 Y 记正面出现次数与反面出现次数之差的绝对值，求随机变量 (X,Y)的分布律.

解 X 的可取值为 0，1，2，3，Y 的可取值为 1，3. 由 $Y=|X-(3-X)|$知：当$X=0,3$ 时，$Y=3$；当$X=1,2$ 时，$Y=1$. 利用二项分布可得

$$P\{X=0,Y=3\}=(1/2)^3=1/8,$$
$$P\{X=1,Y=1\}=C_3^1\times1/2\times(1/2)^2=3/8,$$
$$P\{X=2,Y=1\}=C_3^2\times(1/2)^2\times1/2=3/8,$$
$$P\{X=3,Y=3\}=C_3^3\times(1/2)^3=1/8.$$

所以，(X,Y)的联合分布律为

Y \ X	0	1	2	3
1	0	3/8	3/8	0
3	1/8	0	0	1/8

例 3 已知 X 的概率分布为

$$P\{X=-2\}=P\{X=-1\}=P\{X=1\}=P\{X=2\}=1/4,$$

求:(1) $Y=X^2$ 的分布律; (2) 求(X,Y)的分布律.

解 $Y=X^2$ 的分布律为

Y	1	4
p_k	1/2	1/2

因为$X=-2$,$Y=4$,所以,由事件的等价性,有

$$P\{X=-2,Y=4\}=P\{X=-2\}=1/4.$$

同理 $$P\{X=-1,Y=1\}=P\{X=-1\}=1/4,$$

$$P\{X=1,Y=1\}=P\{X=1\}=1/4,$$

$$P\{X=2,Y=4\}=P\{X=2\}=1/4.$$

所以,(X,Y)的联合分布律为

Y \ X	-2	-1	1	2
1	0	1/4	1/4	0
4	1/4	0	0	1/4

例 4 一箱零件有 10 个,其中有 2 个一级品,7 个二级品,1 个次品. 从中任取 3 个,用X记其中的一级品数,Y记其中的二级品数,求(X,Y)的联合分布律.

解 X的可取值为0,1,2,Y的可取值为0,1,2,3. 由题设知,$2\leqslant X+Y\leqslant 3$,故

$$P\{X=0,Y=0\}=P\{X=0,Y=1\}=0,$$

$$P\{X=1,Y=0\}=P\{X=1,Y=3\}=0,$$

$$P\{X=2,Y=2\}=P\{X=2,Y=3\}=0,$$

$$P\{X=0,Y=2\}=C_2^0C_7^2C_1^1/C_2^0=7/40,$$

$$P\{X=0,Y=3\}=C_2^0C_7^3C_1^0/C_{10}^3=7/24,$$

$$P\{X=1,Y=1\}=C_2^1C_7^1C_1^1/C_{10}^3=7/60,$$

$$P\{X=1,Y=2\}=C_2^1C_7^2C_1^0/C_{10}^3=7/20,$$

$$P\{X=2,Y=0\}=C_2^2C_7^0C_1^1/C_{10}^3=1/120,$$

$$P\{X=2,Y=1\}=C_2^2C_7^1C_1^0/C_{10}^3=7/120.$$

所以，(X,Y)的联合分布律为

$X \backslash Y$	0	1	2	3
0	0	0	7/40	7/24
1	0	7/60	7/20	0
2	1/120	7/120	0	0

例5 一盒内装有大小相同的21个球，分别标有号码1,2,…,21.现从盒中随机取出一球，以$X=0$和$X=1$分别记取得球的号码为偶数和奇数的事件，以$Y=0$和$Y=1$分别记取得球的号码为3的倍数与不是3的倍数的事件，求(X,Y)的联合分布律.

解 (X,Y)的可能取值为(0,0),(0,1),(1,1),(1,0)，要找出它们的积事件.

样本空间基本事件总事件数为21.(0,0)含号码为6,12,18等三个事件，(0,1)含2,4,8,10,14,16,20等七个事件，(1,0)含3,9,15,21等四个事件，(1,1)含1,5,7,11,13,17,19等七个事件.由古典型概率计算公式，求出

$$P(0,0)=1/7,\quad P(0,1)=1/3,$$

$$P(1,0)=4/21,\quad P(1,1)=1/3.$$

所以，(X,Y)的联合分布律为

$Y \backslash X$	0	1
0	1/7	4/21
1	1/3	1/3

例6 将一均匀硬币掷三次，以X记前两次正面出现的次数，以Y记三次中正面出现的次数，求(X,Y)的联合分布律.

解 X可取值为0,1,2，Y可取值为0,1,2,3.显然(0,2),(0,3),(1,0),(1,3),(2,0),(2,1)为不可能事件，概率为零.而

$$P\{X=0,Y=0\}=P\{X=0,Y=1\}=(1/2)^3=1/8,$$

$$P\{X=1,Y=1\}=C_2^1\times(1/2)^2\times1/2=1/4,$$

$$P\{X=1,Y=2\}=C_2^1\times(1/2)^2\times 1/2=1/4,$$
$$P\{X=2,Y=2\}=P\{X=2,Y=3\}=(1/2)^3=1/8,$$
所以,(X,Y)的联合分布律为

X \ Y	0	1	2	3
0	1/8	1/8	0	0
1	0	1/4	1/4	0
2	0	0	1/8	1/8

例7 设离散型随机变量(X,Y)的分布为

$$p_{ij}=(i+j)/30,\ i=0,1,2,3 \text{ 且 } j=0,1,2.$$

求:(1) $P\{X>2,Y\leqslant 2\}$; (2) $P\{X>Y\}$; (3) $P\{X+Y=4\}$.

解 (1) 事件$\{X>0,Y\leqslant 2\}$包含$(3,2)$,$(3,1)$,$(3,0)$,所以

$$P\{X>2,Y\leqslant 2\}=[(3+2)+(3+1)+(3+0)]/30=2/5.$$

(2) 事件$\{X>Y\}$包含$(1,0)$,$(2,1)$,$(2,0)$,$(3,2)$,$(3,1)$,$(3,0)$,所以

$$P\{X>Y\}=(1+3+2+5+4+3)/30=3/5.$$

(3) 事件$\{X+Y=4\}$包含$(2,2)$,$(3,1)$,所以

$$P\{X+Y=4\}=(4+4)/30=4/15.$$

例8 设(X,Y)的分布函数为$F(x,y)$,试用$F(x,y)$表示:

(1) $P\{a\leqslant x\leqslant b,Y<c\}$; (2) $P\{0<Y<b\}$;

(3) $P\{X\geqslant a,Y<b\}$.

解 $P\{a\leqslant x\leqslant b,Y<c\}=F(b,c)-F(a,c)$,

$$P\{0<Y<b\}=F(+\infty,b)-F(+\infty,0),$$
$$P\{X\geqslant a,Y<b\}=1+F(a,b)-F(+\infty,b)-F(a,+\infty).$$

三、二维连续型随机变量(X,Y)的计算通常存在的几个问题

(1) 已知分布形式,求分布的密度函数和分布函数.解决这类问题的方法,一般是依据分布的定义,由题给条件讨论.要注意不同区域上密度函数的不同表示形式,将密度函数与分布函数写成分段函数形式.

(2) 已知函数,讨论其是否是二维随机变量的分布函数或密

度函数.解决这类问题的主要依据是,已知函数是否符合定义,特别要注意验证是否满足分布函数或密度函数的性质.

(3) 已知函数,确定函数中的参数,求分布函数或密度函数,并求事件的概率.解决这类问题应首先依据分布函数或密度函数的性质,然后分区域讨论,正确确定各积分区域的积分限,计算二重积分,依据结果写出分布函数或密度函数.在求给定区域上概率时,关键是将区域用不等式组表出,顺利地将$\iint\limits_G f(x,y)\mathrm{d}x\mathrm{d}y$化为二次积分.要注意积分区域的分块与积分限的配置.

例9 说明函数

$$F(x,y)=\begin{cases}0, & x<0\text{ 或 }y<0\text{ 或 }x+y<1,\\ 1, & \text{其它}\end{cases}$$

不是二维随机变量的分布函数.

解 虽然有

$$F(-\infty,y)=F(x,-\infty)=0,\quad F(+\infty,+\infty)=1,$$

且$F(x,y)$对x与y右连续和单调不减,但是,若取$x_1=y_1=0.1$,$x_1=y_2=1.1$,则有

$$\begin{aligned}&F(x_2,y_2)-F(x_2,y_1)-F(x_1,y_2)+F(x_1,y_1)\\&=1-1-1+0=-1.\end{aligned}$$

这与$P\{x_1<X<x_2,y_1<Y<y_2\}\geqslant 0$矛盾,故$F(x,y)$不是二维随机变量$(X,Y)$的分布函数.

例10 说明函数

$$F(x,y)=\begin{cases}1/2+(1-\mathrm{e}^{-x})(1+\mathrm{e}^{-y}), & x>0,y>0,\\ 1/2, & \text{其它}\end{cases}$$

是否为二维随机变量(X,Y)的分布函数.

解 不是.因为它不满足分布函数的性质

$$\lim_{x\to-\infty,y\to-\infty}F(x,y)=1/2\neq 0.$$

例11 设$g(x)\geqslant 0$,且$\int_0^{+\infty}g(x)\mathrm{d}x=1$,有

$$f(x,y)=\begin{cases}\dfrac{2g(\sqrt{x^2+y^2})}{\pi\sqrt{x^2+y^2}}, & 0\leqslant x,y<+\infty,\\ 0, & \text{其它}.\end{cases}$$

问：$f(x,y)$是否可以作为二维随机变量(X,Y)的概率密度函数？

解 可以.显然，$f(x,y)\geqslant 0$，是不减函数，又

$$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\frac{2g(\sqrt{x^2+y^2})}{\pi\sqrt{x^2+y^2}}\mathrm{d}x\mathrm{d}y$$
$$=\int_0^{\pi/2}\mathrm{d}\theta\int_0^{+\infty}\frac{2}{\pi}g(r)\mathrm{d}r=\int_0^{+\infty}g(r)\mathrm{d}r=1$$

符合概率密度函数的性质，所以，可以作为二维随机变量(X,Y)的概率密度函数.

例12 说明函数

$$f(x,y)=\begin{cases}x^2+y^2, & x^2+y^2<2,\\ 0, & \text{其它}\end{cases}$$

不能作为随机变量(X,Y)的概率密度函数.

解 因为

$$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}f(x,y)\mathrm{d}x\mathrm{d}y=\iint\limits_{x^2+y^2<2}(x^2+y^2)\mathrm{d}x\mathrm{d}y$$
$$=\int_0^{2\pi}\mathrm{d}\theta\int_0^{\sqrt{2}}r^2\mathrm{d}r=2\pi\neq 1$$

不符合概率密度函数性质，所以不能作为(X,Y)的概率密度函数.

例13 求在区域G上服从均匀分布的随机变量(X,Y)的密度函数与分布函数，其中G由直线$x=0,y=0,y=x+1$所围成.

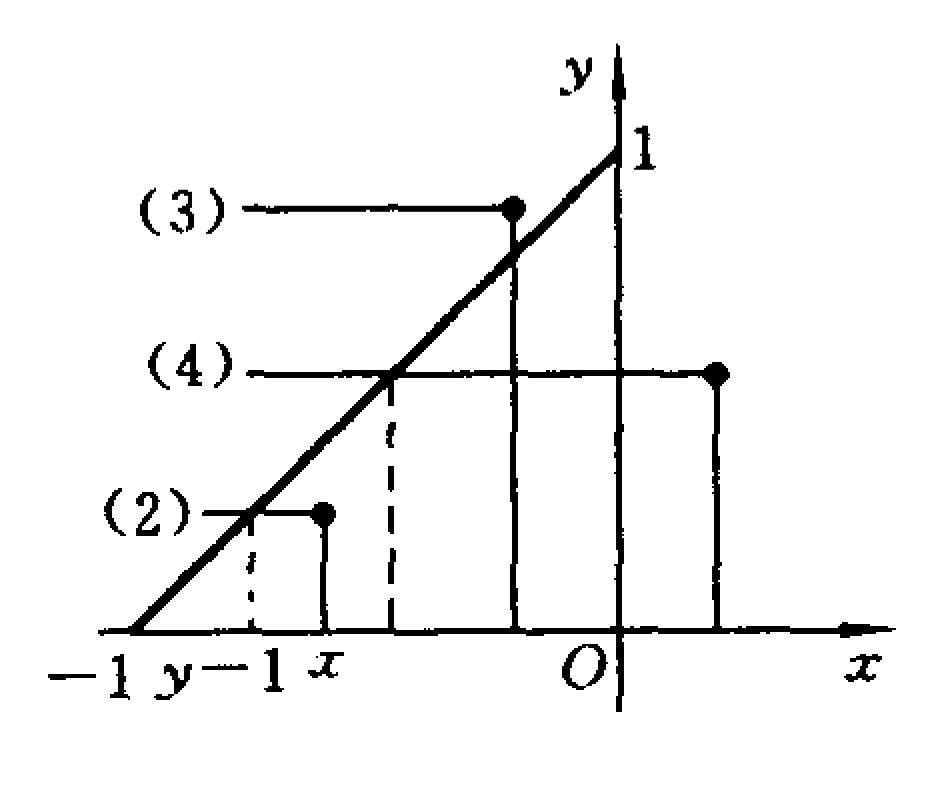


图3.3

解 如图3.3所示，先计算G的面积S_G，$S_G=1/2$，所以(X,Y)的联合密度为

$$f(x,y)=\begin{cases}2, & (x,y)\in G,\\ 0, & \text{其它}.\end{cases}$$

求$F(x,y)$需要分区域进行讨论.

(1) 当 $x<-1$ 或 $y<0$ 时，有

$$f(x,y)=0,\quad F(x,y)=0.$$

(2) 当 $-1\leqslant x<0, 0\leqslant y<x+1$ 时，$f(x,2)=2$，有

$$F(x,y)=\int_{-1}^{y-1}\mathrm{d}x\int_{0}^{x+1}2\mathrm{d}y+\int_{y-1}^{x}\mathrm{d}x\int_{0}^{y}2\mathrm{d}y=(2x-y+2)y.$$

(3) 当 $-1\leqslant x<0, y\geqslant x+1$ 时，有

$$F(x,y)=\int_{-1}^{x}\mathrm{d}x\int_{0}^{y}2\mathrm{d}y=(x+1)^2.$$

(4) 当 $x\geqslant 0, 0\leqslant y<1$ 时，有

$$F(x,y)=\int_{-1}^{y-1}\mathrm{d}x\int_{0}^{x+1}2\mathrm{d}y+\int_{y-1}^{0}\mathrm{d}x\int_{0}^{y}2\mathrm{d}y=(2-y)y.$$

(5) 当 $x\geqslant 0, y\geqslant 1$ 时，有 $F(x,y)=1$.

所以 $F(x,y)$ 是一分段函数

$$F(x,y)=\begin{cases}0, & x<-1 \text{ 或 } y<0,\\ (2x-y+2)y, & -1\leqslant x<0, 0\leqslant y<x+1,\\ (x+1)^2, & -1\leqslant x<0, y\geqslant x+1,\\ (2-y)y, & x\geqslant 0, 0\leqslant y<1,\\ 1, & x\geqslant 0, y\geqslant 1.\end{cases}$$

例14 随机变量 (X,Y) 在区域 G：$a\leqslant x\leqslant b, c\leqslant y\leqslant d$ 内服从均匀分布，求 $f(x,y)$ 与 $F(x,y)$.

解 设 (X,Y) 的概率密度为

$$f(x,y)=\begin{cases}A, & a\leqslant x\leqslant b, c\leqslant y\leqslant d,\\ 0, & \text{其它},\end{cases}$$

则

$$\begin{aligned}1&=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}f(x,y)\mathrm{d}x\mathrm{d}y=\int_{a}^{b}\mathrm{d}x\int_{c}^{d}A\mathrm{d}y\\ &=A(b-a)(d-c),\end{aligned}$$

所以

$$A=1/[(b-a)(d-c)],$$

$$f(x,y)=\begin{cases}1/[(b-a)(d-c)], & a\leqslant x\leqslant b, c\leqslant y\leqslant d,\\ 0, & \text{其它}.\end{cases}$$

同上题方法，分区域讨论(见图 3.4)，得

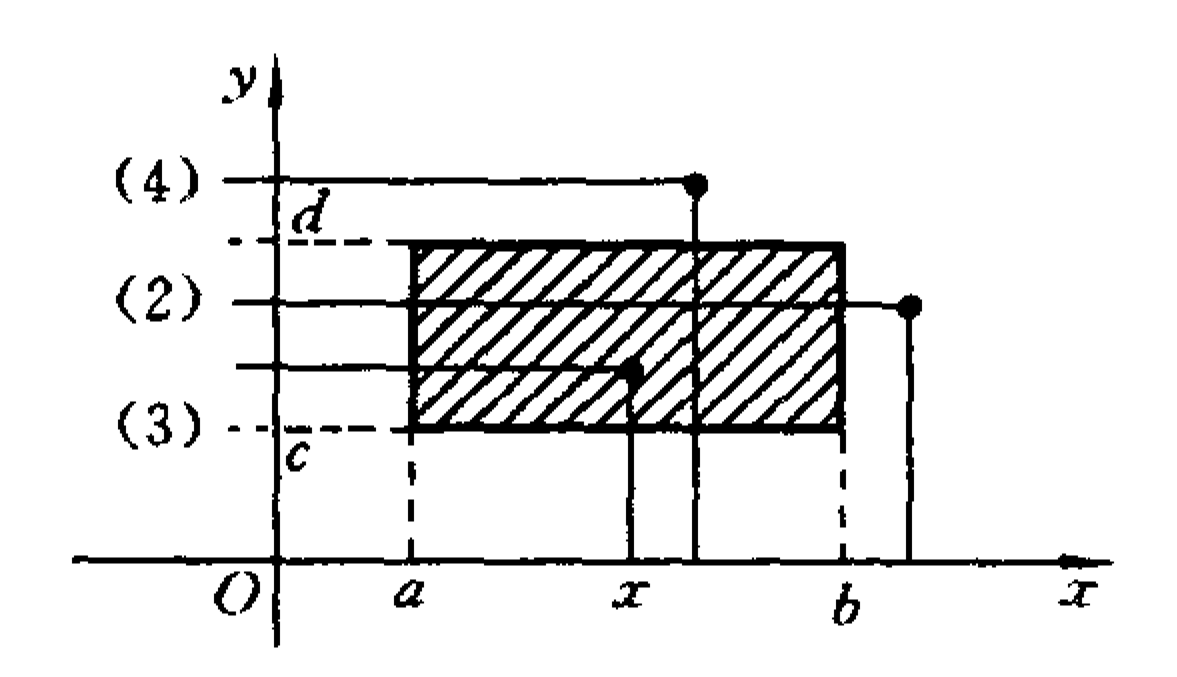


图 3.4

$$F(x,y)=\begin{cases}0, & x<a,y<c,\\ \dfrac{(x-a)(y-c)}{(b-a)(d-c)}, & a\leqslant x<b,c\leqslant y<d,\\ (y-c)/(d-c), & x\geqslant b,c\leqslant y<d,\\ (x-a)/(b-a), & a\leqslant x<b,y\geqslant d,\\ 1, & x\geqslant b,y\geqslant d.\end{cases}$$

例 15 设二维随机变量$(X,Y)\sim N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$，其概率密度为

$$f(x,y)=\frac{1}{2\pi\sqrt{3}}e^{-(4x^2+2xy+y^2-8x-2y+4)/6},\quad -\infty<x,y<+\infty,$$

求参数$\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho$的值.

解 因为

$$\frac{4x^2+2xy+y^2-8x-2y+4}{6}=\frac{4(x-1)^2+2(x-1)y+y^2}{2(\sqrt{3})^2},$$

所以，与二维正态分布的密度函数比较，得

$$\mu_1=1,\quad \mu_2=0,\quad \sigma_1^2=1,\quad \sigma_2^2=4,\quad \rho=-1/2.$$

故知$(X,Y)\sim N(1,0;1,4;-1/2)$.

例 16 设二维随机变量(X,Y)的概率密度为

$$f(x,y)=Ae^{-(x^2+y^2)/200},\quad -\infty<x,y<+\infty,$$

求：(1) 系数A的值； (2) 确定(X,Y)的分布及分布的参数；

(3) $P\{X>Y\}$.

解 (1)$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}Ae^{-(x^2+y^2)/200}dxdy$（极坐标代换）

$$=\int_0^{2\pi}\mathrm{d}\theta\int_0^{+\infty}Ae^{r^2/200}r\mathrm{d}r=200\pi A=1,$$

得 $A=1/(200\pi)$.

(2) $f(x,y)=\frac{1}{200\pi}^{e-(x^2+y^2)/200}=\frac{1}{2\pi\times 100}e^{(x^2/100+y^2/100)/2}$,

所以,$(X,Y)\sim N(0,0,10^2,10^2,0)$,是二维正态分布.

(3) 由对称性,$P\{X>Y\}=P\{X<Y\}=1/2$.

例 17 设随机变量(X,Y)的分布函数为

$$F(x,y)=\begin{cases}c-3^{-x}-3^{-y}+3^{-x-y}, & x\geqslant 0,y\geqslant 0,\\ 0, & \text{其它},\end{cases}$$

求:(1)常数c; (2)概率密度函数$f(x,y)$.

解 (1)由$1=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}(c-3^{-x}-3^{-y}+3^{-x-y})\mathrm{d}x\mathrm{d}y=c$,得$c=1$.

(2) $f(x,y)=\frac{\partial^2 F(x,y)}{\partial x\partial y}=\frac{\partial}{\partial y}(3^{-x}\ln 3-3^{-x-y}\ln 3)$

$$=3^{-x-y}(\ln 3)^2,\quad x\geqslant 0,y\geqslant 0.$$

故
$$f(x,y)=\begin{cases}3^{-x-y}(\ln 3)^2, & x\geqslant 0,y\geqslant 0,\\ 0, & \text{其它}.\end{cases}$$

例 18 设随机变量(X,Y)的概率密度为

$$f(x,y)=\begin{cases}A(R-\sqrt{x^2+y^2}), & x^2+y^2\leqslant R^2,\\ 0, & \text{其它},\end{cases}$$

求:(1) 系数 A 的值;

(2) 概率 $P\{(X,Y)\in x^2+y^2\leqslant r^2\}$ $(r\leqslant R)$.

解 需利用极坐标代换

(1) $1=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}f(x,y)\mathrm{d}x\mathrm{d}y$

$$=\iint_{x^2+y^2\leqslant R^2}A(R-\sqrt{x^2+y^2})\mathrm{d}x\mathrm{d}y$$

$$=A\int_0^{2\pi}\mathrm{d}\theta\int_0^R(R-r)r\mathrm{d}r=A\pi R^3/3,$$

得
$$A=3/(\pi R^3).$$

(2) $P\{(X,Y)\in x^2+y^2\leqslant r^2\}=\frac{3}{\pi R^3}\int_0^{2\pi}\mathrm{d}\theta\int_0^r(R-r)r\mathrm{d}r$

$$=\frac{3r^2}{R^3}\left(1-\frac{2r}{3R}\right).$$

例 19 设随机变量 (X,Y) 的概率密度为

$$f(x,y)=\begin{cases}k(6-x-y), & 0<x<2,2<y<4,\\ 0, & \text{其它},\end{cases}$$

求：(1) 常数 k； (2) $P\{X<1,Y<3\}$；

(3) $P\{X+Y\leqslant 4\}$（见图 3.5）.

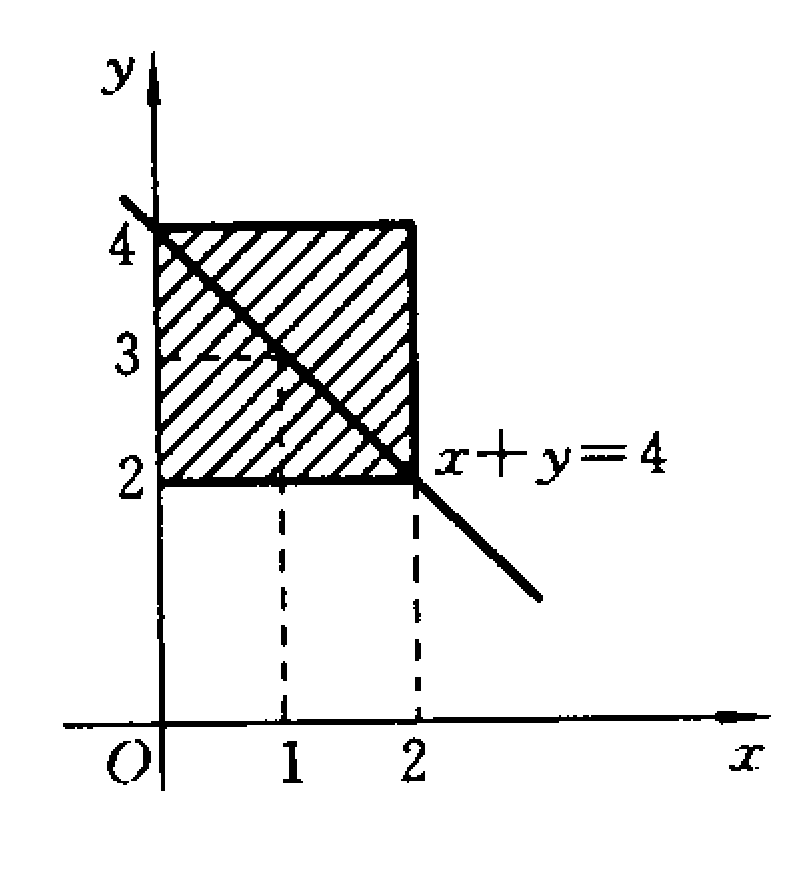


图 3.5

解 (1) $\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}f(x,y)\mathrm{d}x\mathrm{d}y$

$$=\int_2^4\mathrm{d}y\int_0^2 k(6-x-y)\mathrm{d}x$$

$$=k\int_2^4(10-2y)\mathrm{d}y=8k,$$

故 $k=1/8$.

(2) $P\{X<1,Y<3\}=\int_2^3\mathrm{d}y\int_0^1\frac{1}{8}(6-x-y)\mathrm{d}y$

$$=\int_2^3\frac{1}{8}\left(\frac{11}{2}-y\right)\mathrm{d}y=\frac{3}{8}.$$

(3) $P\{X+Y\leqslant 4\}=\frac{1}{8}\int_0^2\mathrm{d}x\int_2^{4-x}(6-x-y)\mathrm{d}y$

$$=\frac{1}{8}\int_0^2\left(6-4x+\frac{x^2}{2}\right)\mathrm{d}x=\frac{2}{3}.$$

例 20 设二维随机变量 (X,Y) 的概率密度为

$$f(x,y)=\begin{cases}6\mathrm{e}^{-(2x+3y)}, & x>0,y>0,\\ 0, & \text{其它},\end{cases}$$

求：(1) $F(x,y)$；

(2) $P\{2X+3Y\leqslant 6\}$.

解 (1) 分区域讨论（见图 3.6）.

当 $x\leqslant 0,y\leqslant 0$ 时，$F(x,y)=0$.

当 $x>0,y>0$ 时，

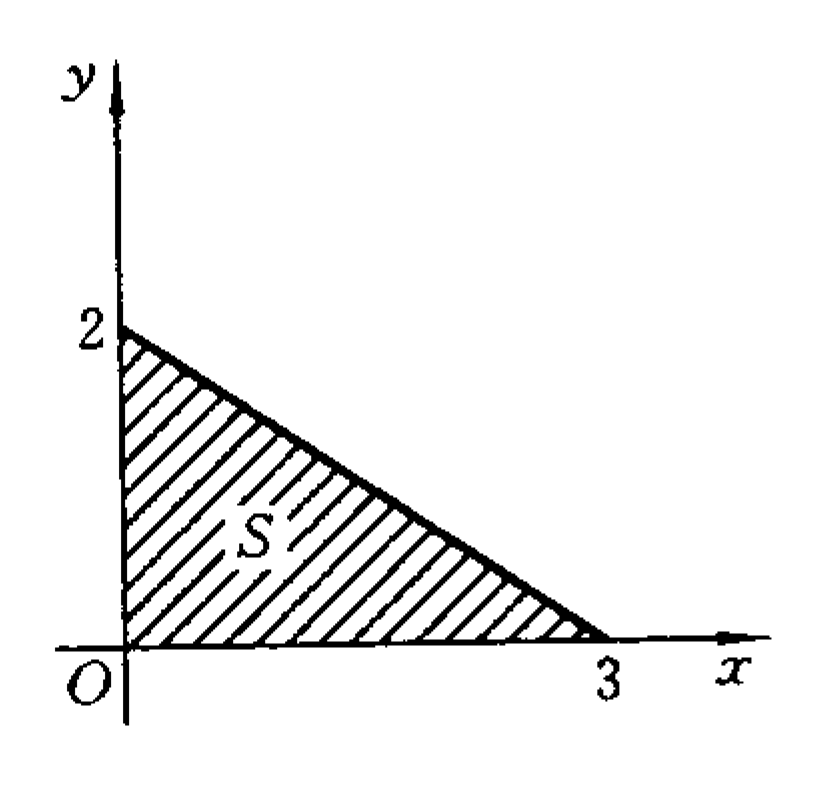


图 3.6

$$F(x,y)=\int_0^x \mathrm{d}y\int_0^y 6\mathrm{e}^{-(2x+3y)}\mathrm{d}x=(1-\mathrm{e}^{-2x})(1-\mathrm{e}^{-3y}).$$

即
$$F(x,y)=\begin{cases}(1-\mathrm{e}^{-2x})(1-\mathrm{e}^{-3y}), & x>0,y>0,\\ 0, & \text{其它}.\end{cases}$$

(2) $P\{2X+3Y\leqslant 6\}$

$$=\iint\limits_{2x+3y\leqslant 6} 6\mathrm{e}^{-(2x+3y)}\mathrm{d}x\mathrm{d}y=\int_0^3 \mathrm{d}x\int_0^{2(3-x)/3}\mathrm{e}^{-(2x+3y)}\mathrm{d}y$$

$$=1-7\mathrm{e}^{-6}=0.9826.$$

例21 已知随机变量(X,Y)的概率密度为

$$f(x,y)=\begin{cases}1/2, & 0\leqslant x\leqslant 1,0\leqslant y\leqslant 2,\\ 0, & \text{其它},\end{cases}$$

求:X与Y中至少有一个小于1/2的概率.

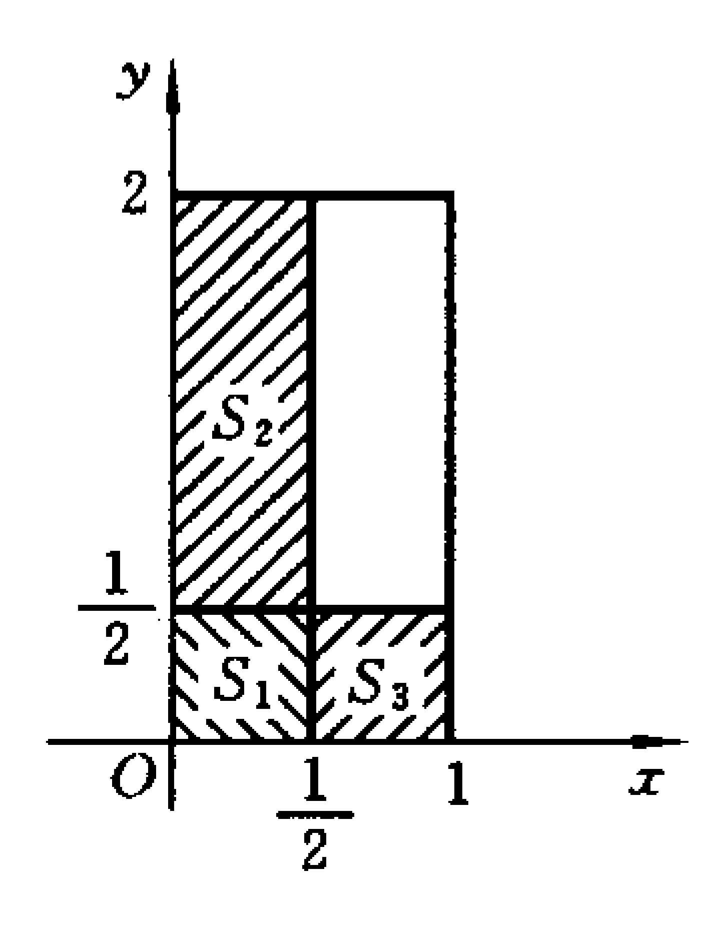


图3.7

解 事件$\{X$与Y至少有一个小于$1/2\}$等价于随机点(X,Y)充满图3.7中$S_1\cup S_2\cup S_3$,所以

$$p=\iint\limits_{S_1+S_2+S_3} f(x,y)\mathrm{d}x\mathrm{d}y$$

$$=\int_0^{1/2}\mathrm{d}x\int_0^{1/2}\frac{1}{2}\mathrm{d}y+\int_0^{1/2}\mathrm{d}x\int_{1/2}^{2}\frac{1}{2}\mathrm{d}y+\int_{1/2}^{1}\mathrm{d}x\int_0^{1/2}\frac{1}{2}\mathrm{d}y$$

$$=(1/2\times 1/2+1/2\times 3/2+1/2\times 1/2)/2=5/8.$$

也可直接用几何型概率求解(记S为边长分别为1和2的矩形面积),则

$$p=(S_1+S_2+S_3)/S$$

$$=(1/2\times 1/2+1/2\times 3/2+1/2\times 1/2)/2=5/8.$$

例22 设随机变量(X,Y,Z)的密度函数为

$$f(x,y,z)=\begin{cases}\mathrm{e}^{-(x+y+z)}, & x>0,y>0,z>0,\\ 0, & \text{其它},\end{cases}$$

求概率$P\{X<Y<Z\}$.

解 即求随机点(X,Y,Z)落入区域G:$\{0<x<+\infty,x<y<$

$+\infty, y<z<+\infty\}$内的概率，所以

$$P\{X<Y<Z\}=\iiint_G f(x,y,z)\mathrm{d}x\mathrm{d}y\mathrm{d}z$$

$$=\int_0^{+\infty}\mathrm{d}x\int_x^{+\infty}\mathrm{d}y\int_y^{+\infty}\mathrm{e}^{-(x+y+z)}\mathrm{d}z$$

$$=\int_0^{+\infty}\mathrm{e}^{-x}\left[\int_x^{+\infty}\mathrm{e}^{-y}\left(\int_y^{+\infty}\mathrm{e}^{-z}\mathrm{d}z\right)\mathrm{d}y\right]\mathrm{d}x=\frac{1}{6}.$$

例 23 设二维随机变量(X,Y)的概率密度为

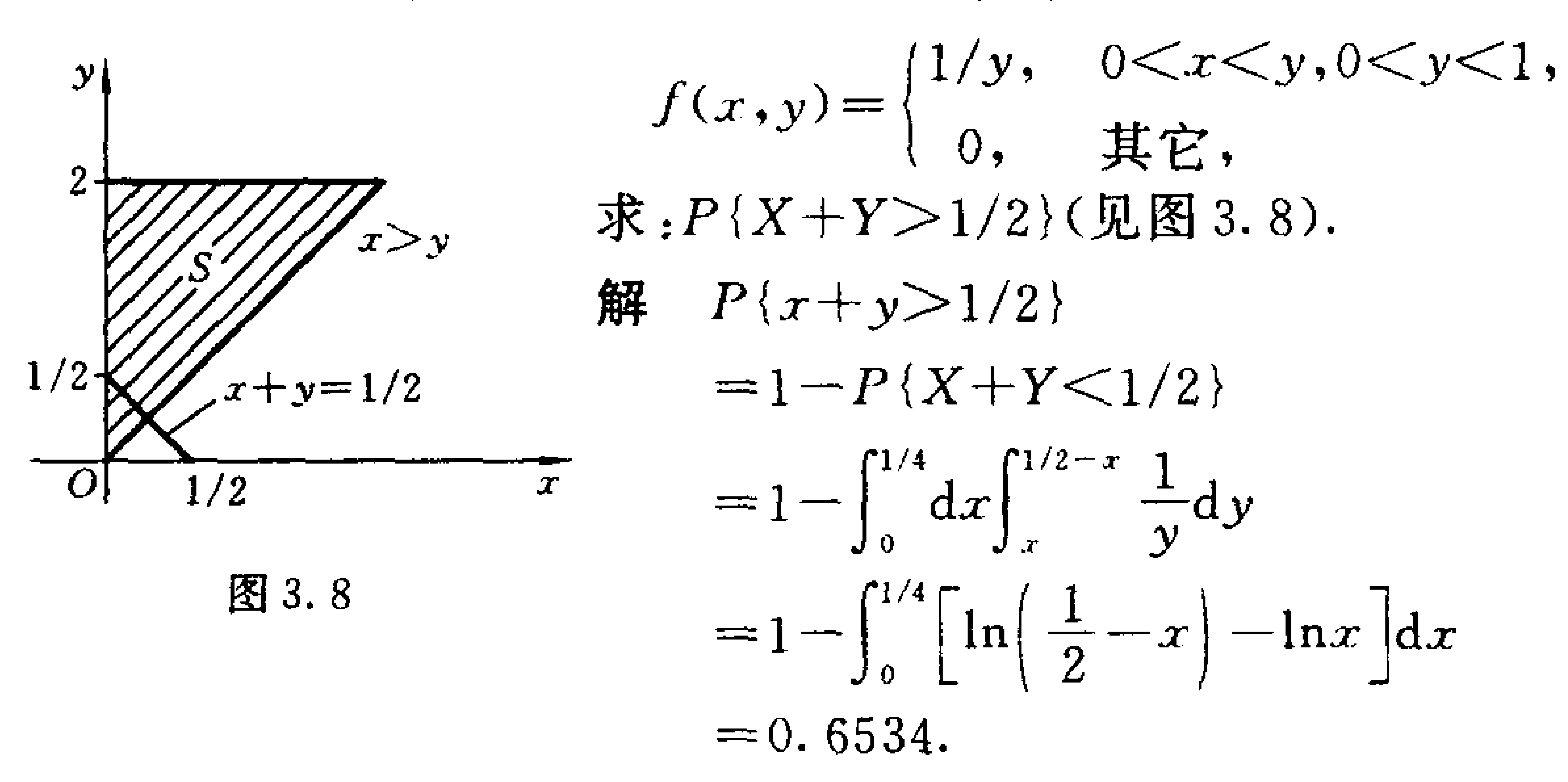


图 3.8

$$f(x,y)=\begin{cases}1/y, & 0<x<y, 0<y<1,\\ 0, & \text{其它},\end{cases}$$

求：$P\{X+Y>1/2\}$（见图 3.8）.

解 $P\{x+y>1/2\}$

$$=1-P\{X+Y<1/2\}$$

$$=1-\int_0^{1/4}\mathrm{d}x\int_x^{1/2-x}\frac{1}{y}\mathrm{d}y$$

$$=1-\int_0^{1/4}\left[\ln\left(\frac{1}{2}-x\right)-\ln x\right]\mathrm{d}x$$

$$=0.6534.$$

第二节 二维随机变量的边缘分布与条件分布

主要内容

一、二维随机变量的边缘分布

1. 边缘分布函数

组成二维随机变量(X,Y)的随机变量X,Y各自的分布函数$F_X(x),F_Y(y)$称为二维随机变量(X,Y)关于X和关于Y的边缘分布函数.

边缘分布函数可以由(X,Y)的分布函数$F(x,y)$确定，即

$$F_X(x)=P\{X\leqslant x,Y<+\infty\}=\lim_{y\to+\infty}F(x,y)=F(x,+\infty),$$

$$F_Y(y)=P\{X<+\infty,Y\leqslant y\}=\lim_{x\to+\infty}F(x,y)=F(+\infty,y).$$

2. 边缘分布律

设二维离散型随机变量(X,Y)的联合分布律为

$$P\{X=x_i,Y=y_j\}=p_{ij},\ i,j=1,2,\cdots,$$

则X,Y的边缘分布律为

$$P\{X=x_i\}=p_{i\cdot}=\sum_{j=1}^{\infty}p_{ij},\quad i=1,2,\cdots,$$

$$P\{Y=y_j\}=p_{\cdot j}=\sum_{i=1}^{\infty}p_{ij},\quad j=1,2,\cdots.$$

3. 边缘概率密度

设二维连续型随机变量(X,Y)的联合概率密度为$f(x,y)$，则关于X及Y的边缘概率密度为

$$f_X(x)=\int_{-\infty}^{+\infty}f(x,y)\mathrm{d}y,\quad f_Y(y)=\int_{-\infty}^{+\infty}f(x,y)\mathrm{d}x.$$

二、二维随机变量(X,Y)的条件分布

1. 条件分布律

设二维离散型随机变量(X,Y)的联合分布律为

$$P\{X=x_i,Y=y_j\}=p_{ij},\quad i,j=1,2,\cdots,$$

关于X和Y的边缘分布分别为

$$P\{X=x_i\}=p_{i\cdot}\quad 和\quad P\{Y=y_j\}=p_{\cdot j},$$

则对于固定的j，若$P\{Y=y_j\}=p_{\cdot j}>0$，称

$$P\{x=X|Y=y_j\}=p_{ij}/p_{\cdot j},\quad i=1,2,\cdots$$

为在条件$Y=y_j$下，随机变量X的条件分布律.

对于固定的i，若$P\{X=x_i\}=p_{i\cdot}>0$，称

$$P\{Y=y|X=x_i\}=p_{ij}/p_{i\cdot},\quad j=1,2,\cdots$$

为在条件$X=x_i$下，随机变量Y的条件分布律.

2. 条件分布函数

给定y，设对于固定的任意$\varepsilon>0$，$P\{y-\varepsilon<Y\leqslant y+\varepsilon\}>0$，且对

任意实数 x，极限

$$\lim_{\varepsilon \to 0^+} P\{X \leqslant x \mid y-\varepsilon < Y \leqslant y+\varepsilon\}$$
$$= \lim_{\varepsilon \to 0^+} P\{X \leqslant x, y-\varepsilon < Y \leqslant y+\varepsilon\} / P\{y-\varepsilon < Y \leqslant y+\varepsilon\}$$

存在，则称此极限为在条件 $Y=y$ 下 X 的条件分布函数，记为

$$F_{X|Y}(x \mid y).$$

设 (X,Y) 的联合分布函数为 $F(x,y)$，联合概率密度为 $f(x,y)$. 若在点 (x,y)，$f(x,y)$ 连续，边缘概率密度 $f_Y(y)$ 连续，且 $f_Y(y)>0$，则

$$F_{X|Y}(x \mid y) = \frac{\int_{-\infty}^{x} f(u,y)\mathrm{d}u}{f_Y(y)} = \int_{-\infty}^{x} \frac{f(u,y)}{f_Y(y)}\mathrm{d}u.$$

称 $f_{X|Y}(x \mid y) = \dfrac{f(x,y)}{f_Y(y)}$ 为在条件 $Y=y$ 下 X 的条件概率密度.

类似地定义

$$F_{Y|X}(y \mid x) = \frac{\int_{-\infty}^{y} f(x,v)\mathrm{d}v}{f_X(x)} = \int_{-\infty}^{y} \frac{f(x,v)}{f_X(x)}\mathrm{d}v,$$

$$f_{Y|X}(y \mid x) = \frac{f(x,y)}{f_X(x)}.$$

二维随机变量 (X,Y) 的联合分布唯一确定边缘分布，也唯一确定条件分布. 反之却不一定成立.

疑 难 解 析

1. 二维随机变量 (X,Y) 的联合分布、边缘分布和条件分布之间存在什么样的关系？

答　由定义知，(X,Y) 的联合分布唯一确定关于 X 和关于 Y 的边缘分布，也唯一确定条件分布. 反之，边缘分布与条件分布却不一定能唯一确定联合分布. 但由

$$f(x,y) = f_X(x) f_{Y|X}(y \mid x) = f_Y(y) f_{X|Y}(x \mid y)$$

知，一个条件分布和它对应的边缘分布能唯一确定一个联合分布.

例如，二维正态分布$(X,Y)\sim N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$的边缘分布是$X\sim N(\mu_1,\sigma_1^2)$和$Y\sim N(\mu_2,\sigma_2^2)$，很明显与$\rho$无关；而$(X,Y)$的边缘分布$X,Y$，当$\rho$不同时却可以得到不同的联合分布

$$N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho).$$

学完下一节后我们将知道，当组成(X,Y)的X,Y相互独立时，有$P\{X\leqslant x,Y\leqslant y\}=P\{X\leqslant x\}P\{Y\leqslant y\}$，即$F(x,y)=F_X(x)F_Y(y)$. 于是知，$X,Y$相互独立时，边缘分布能唯一确定联合分布，从而知条件分布也能唯一确定联合分布.

2. 二维随机变量的边缘分布与一维随机变量的分布有什么联系与区别?

答 从某种意义上讲，二维随机变量的每个边缘分布是一维随机变量的分布. 如，二维正态分布$(X,Y)\sim N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$的边缘分布$X\sim N(\mu_1,\sigma_1^2)$，$Y\sim N(\mu_2,\sigma_2^2)$具备一维分布的性质，所以说，边缘分布与一维分布有联系.

但是从严格意义上讲，二维随机变量的边缘分布是定义在$\mathbf{R}^2$平面上的，而一维随机变量的分布是定义在实轴上的，两者的定义域不同. 如(X,Y)的边缘分布$F_X(x)=P\{X\leqslant x,Y<+\infty\}$表示随机点$(X,Y)$落在区域$\{-\infty<X<x,-\infty<Y<+\infty\}$内的概率，而$F(x)=P\{X\leqslant x\}$表示随机点$X$落在区间$(-\infty,x]$上的概率. 两者是有区别的.

3. 为什么不能用条件概率定义直接定义连续型随机变量的条件分布?

答 在第一章中得知，$P(A|B)=P(AB)/P(B)$当$P(B)>0$时成立，在离散型随机变量时，可以借此定义，用条件概率定义直接定义条件分布

$$P\{X=x_i|Y=y_j\}=P\{X=x_i,Y=y_j\}/P\{Y=y_j\}=p_{ij}/p_{\cdot j}.$$

但是，在连续型随机变量的情形，因为在任一点(x,y)，概率

$$P\{X=x,Y=y\}=0,\quad P\{X=x\}=P\{Y=y\}=0,$$

所以,不能用条件概率定义来定义条件分布.要定义一个区间

$$\{x-\varepsilon<X\leqslant x+\varepsilon\} \quad 或 \quad \{y-\varepsilon<Y\leqslant y+\varepsilon\},$$

使 $P\{x-\varepsilon<X\leqslant x+\varepsilon\}>0$ 或 $P\{y-\varepsilon<Y\leqslant y+\varepsilon\}>0$,才可以作为分式的分母,然后用 $\varepsilon\to 0$ 的极限得出条件分布的定义

$$F_{X|Y}(x|y)=P\{X\leqslant x|Y=y\}=\lim_{\varepsilon\to 0}\frac{P\{X\leqslant x,y-\varepsilon<Y\leqslant y+\varepsilon\}}{P\{y-\varepsilon<Y\leqslant y+\varepsilon\}}.$$

4. 两个正态随机变量的联合分布一定是正态随机变量吗?

答 不一定.一个二维正态随机变量的两个边缘分布也是正态随机变量,但是,当两个边缘分布都是正态随机变量时,其联合分布未必是正态随机变量.例如,设 $X\sim N(0,1)$,$Y\sim N(0,1)$,而 (X,Y) 的联合分布密度函数为

$$f(x,y)=\frac{1}{2\pi}\exp\left[-\frac{1}{2}(x^2+y^2)\right]\cdot(1+\sin x\sin y)$$

时,(X,Y) 就不是正态随机变量.

同时指出,由两个随机变量 X,Y 的联合分布密度 $f(x,y)$ 容易求出 X,Y 各自的边缘分布密度 $f_X(x)$ 和 $f_Y(y)$,但是,已知 $f_X(x)$ 和 $f_Y(y)$ 时,未必能求出联合分布密度 $f(x,y)$.

方法、技巧与典型例题分析

一、已知联合分布求边缘分布问题

首先要区别离散型和连续型的不同情形.对于离散型的情形,只要将联合分布律的各横行或各纵列元素分别相加,填在联合分布律的边缘上(这也就是边缘分布这一名词的来历)即得.对于连续型的情形,一般利用定义用积分求出.但要注意,在 $f(x,y)$ 为分段函数时,积分也要分段求出;特别是当 $f(x,y)$ 仅在某个区域不为零时,要作出 G 的图形,用"穿线法"确定积分限,使二重积分计算得出正确结果.

二、连续型随机变量的条件分布的求法

在计算条件分布时,离散型的情形可以由公式直接得出,比较

简单，连续型随机变量的情形则较为复杂. 如求 $f_{X|Y}(x|y)$时，首先要排除 $f_Y(y)=0$ 的区域，即仅对 $f_Y(y)>0$ 才有定义. 其次，要考察$f(x,y)$的值，在$f(x,y)$为零的区域，$f_{X|Y}(x|y)=0$；若$f(x,y)$是分段函数，$f_{X|Y}(x|y)$也要分段表示.

例 1 设二维随机变量(X,Y)联合分布律为

Y \ X	1	2	3
1	0	1/6	1/12
2	1/5	1/9	0
3	2/15	1/4	1/18

求：(1) $P\{X=x_i\}, P\{Y=y_j\}$； (2) $Y=1$ 下 X 的条件分布律.

解 (1) 将纵列或横行各数分别相加，得 $P\{X=x_i\}, P\{Y=y_j\}$，其分布律分别为

X	1	2	3
p_k	1/3	19/36	5/36

Y	1	2	3
p_k	1/4	14/45	79/180

(2) 因为

$$P\{Y=1\}=1/4, \quad P\{X=x_i|Y=1\}=p_{i1}/(1/4),$$

故$Y=1$ 下 X 的条件分布律为

$$P\{X=1|Y=1\}=0/(1/4)=0,$$
$$P\{X=2|Y=1\}=(1/6)/(1/4)=2/3,$$
$$P\{X=3|Y=1\}=(1/12)/(1/4)=1/3.$$

例 2 (X,Y)的联合分布律为

X \ Y	1	2	3	4
0	0	1/16	0	3/16
1	1/8	1/8	1/16	0
2	1/16	1/16	3/16	1/8

求：(1) X 和Y 的边缘分布律； (2) $X=1$ 下 Y 的条件分布律.

解 (1) 将横行或纵列各数分别相加，得 X 和 Y 的边缘分布律为

X	1	2	3
p_k	1/4	5/16	7/16

Y	1	2	3	4
p_k	3/16	1/4	1/4	5/16

(2) 因为 $P\{X=1\}=5/16$，

由 $P\{Y=y_j|X=1\}=p_{1j}/p_{1\cdot}$，

得$X=1$下Y的条件分布律为

$$P\{Y=1|X=1\}=(1/8)/(5/16)=2/5,$$
$$P\{Y=2|X=1\}=(1/8)/(5/16)=2/5,$$
$$P\{Y=3|X=1\}=(1/16)/(5/16)=1/5,$$
$$P\{Y=4|X=1\}=0.$$

例3 将某医药公司9月份和8月份收到的青霉素针剂订单分别记为X和Y.据以往的资料知,X和Y的联合分布律为

X \ Y	51	52	53	54	55
51	0.06	0.05	0.05	0.01	0.01
52	0.07	0.05	0.01	0.01	0.01
53	0.05	0.10	0.10	0.05	0.05
54	0.05	0.02	0.01	0.01	0.03
55	0.05	0.06	0.05	0.01	0.03

求:(1) 边缘分布律;(2) 8月份的订单数为51时,9月份订单数的条件分布律.

解 (1)将横行或纵列各元素分别相加,即得关于X和Y的边缘分布律为

k	51	52	53	54	55
$p_{k\cdot}$	0.18	0.15	0.35	0.12	0.20
$p_{\cdot k}$	0.28	0.28	0.22	0.09	0.13

(2) 由$P\{Y=51\}=0.28$,对照$X=k$值,得条件分布律为

k	51	52	53	54	55
$P\{X=k\|Y=51\}$	6/28	7/28	5/28	5/28	5/28

例4 以X记某医院一天出生婴儿的个数，以Y记其中男婴的个数，设X和Y的联合分布律为

$$P\{X=n,Y=m\}=\frac{e^{-14}\times 7.14^m\times 6.86^{n-m}}{m!\ (n-m)!},$$

$$m=0,1,\cdots,n;\ n=0,1,2,\cdots,$$

求：(1) 边缘分布律；(2) 条件分布律；(3) 当$X=20$时，Y的条件分布律.

解 由题给条件，得

(1) $$P\{X=n\}=\sum_{m=0}^{n}P\{X=n,Y=m\}$$

$$=\sum_{m=0}^{n}\frac{e^{-14}\times 7.14^m\times 6.86^{n-m}}{m!(n-m)!}$$

$$=\frac{e^{-14}}{n!}\sum_{m=0}^{n}\frac{n!\times 7.14^m\times 6.86^{n-m}}{m!(n-m)!}$$

$$=\frac{e^{-14}}{n!}\sum_{m=0}^{n}C_n^m\times 7.14^m\times 6.86^{n-m}$$

$$=\frac{e^{-14}}{n!}14^n,\ n=0,1,2,\cdots,$$

$$P\{Y=m\}=\sum_{n=0}^{\infty}e^{-14}\frac{7.14^m\times 6.86^{n-m}}{m!(n-m)!}$$

$$=\frac{e^{-14}}{m!}\times 7.14^m\sum_{k=0}^{\infty}\frac{6.86^k}{k!}$$

$$=\frac{e^{-14}}{m!}\times 7.14^m\times e^{6.86}=e^{-7.14}\times\frac{7.14^m}{m!},$$

$$m=0,1,\cdots,n\ (k=n-m).$$

(2) $$P\{X=n|Y=m\}=P\{X=n,Y=m\}/P\{Y=m\}$$

$$=6.86^{n-m}/[(n-m)!e^{6.86}],\ n=m,m+1,\cdots,$$

$$P\{Y=m|X=n\}=P\{X=n,Y=m\}/P\{X=n\}$$

$$=C_n^m\times(7.14/14)^m\times(6.86/14)^{n-m},\ m=0,1,2,\cdots,n,$$

(3) $$P\{Y=m|X=20\}=C_{20}^m\times 0.51^m\times 0.49^{20-m},$$

$$m=0,1,\cdots,n.$$

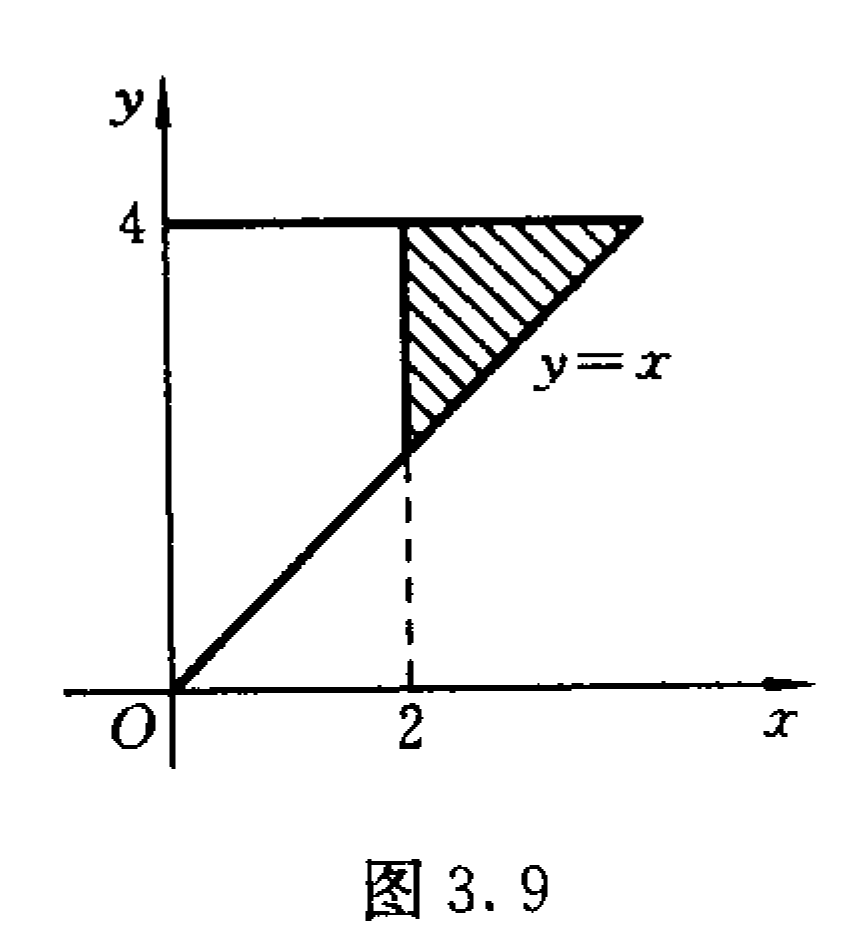


图 3.9

例 5 随机变量(X,Y)的概率密度为

$$f(x,y)=\begin{cases} e^{-y}, & x>0, y>x, \\ 0, & \text{其它}, \end{cases}$$

求:$P\{X>2|Y<4\}$.

解 如图 3.9 所示. 因为

$$f_Y(y)=\begin{cases} \int_0^y e^{-y}dx = ye^{-y}, & y>0, \\ 0, & y\leqslant 0, \end{cases}$$

而 $$P\{X>0,Y>4\}=\iint\limits_G f(x,y)dxdy=\int_2^4 dy\int_2^y e^{-y}dx$$

$$=\int_2^4 (y-2)e^{-y}dy=e^{-2}-3e^{-4}.$$

$$P\{y<4\}=\int_0^4 ye^{-y}dy=1-5e^{-4},$$

所以 $$P\{X>0|Y<4\}=(e^{-2}-3e^{-4})/(1-5e^{-4}).$$

例 6 雷达的圆形屏幕半径为R,如图 3.10 所示,设目标出现点(X,Y)在屏幕上均匀分布,概率密度为

$$f(x,y)=\begin{cases} 1/(\pi R^2), & x^2+y^2\leqslant R^2, \\ 0, & \text{其它}, \end{cases}$$

求:(1) $f_X(x)$,$f_Y(y)$;

(2) $f_{X|Y}(x|y)$,$f_{Y|X}(y|x)$.

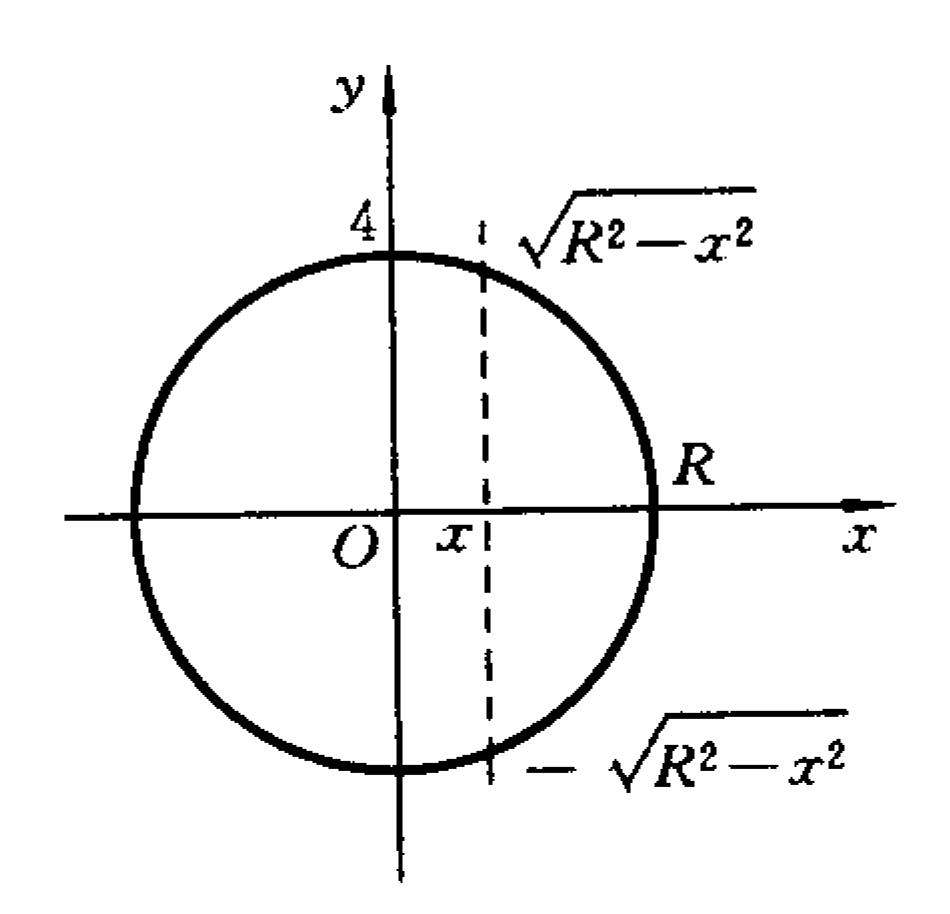


图 3.10

解 (1) 仅当$-R\leqslant x\leqslant R$,

$$-\sqrt{R^2-x^2}\leqslant y\leqslant \sqrt{R^2-x^2}$$

时,$f(x,y)\neq 0$,所以

$$f_X(x)=\begin{cases} \int_{-\sqrt{R^2-x^2}}^{\sqrt{R^2-x^2}} \frac{1}{\pi R^2}dy=\frac{2}{\pi R^2}\sqrt{R^2-x^2}, & -R\leqslant x\leqslant R, \\ 0, & \text{其它}. \end{cases}$$

类似地, $$f_Y(y)=\begin{cases} \frac{2}{\pi R^2}\sqrt{R^2-y^2}, & -R\leqslant y\leqslant R, \\ 0, & \text{其它}. \end{cases}$$

(2) 由 $f_{X|Y}(x|y)=f(x,y)/f_Y(y)$，得

$$f_{X|Y}(x|y)=\begin{cases}1/(2\sqrt{R^2-x^2}), & -\sqrt{R^2-y^2}\leqslant x\leqslant\sqrt{R^2-y^2},\\ 0, & \text{其它}.\end{cases}$$

类似地，

$$f_{Y|X}(y|x)=\begin{cases}1/(2\sqrt{R^2-y^2}), & -\sqrt{R^2-x^2}\leqslant y\leqslant\sqrt{R^2-x^2},\\ 0, & \text{其它}.\end{cases}$$

可以看出，当联合分布是均匀分布时，边缘分布不一定是均匀分布.

例 7 设随机变量 $X\sim N(m,r^2)$，在 $X=x$ 的条件下 Y 的条件分布为 $N(x,\sigma^2)$，求 Y 的概率密度.

解 已知 $f(x,y)=f_{X|Y}(y|x)f_X(x)$，而

$$f_{Y|X}(y|x)=\frac{1}{\sqrt{2\pi}\sigma}\mathrm{e}^{-(y-x)^2/(2\sigma^2)},\quad f_X(x)=\frac{1}{\sqrt{2\pi}r}\mathrm{e}^{-(x-m)^2/(2r^2)},$$

所以

$$\begin{aligned}f_Y(y)&=\int_{-\infty}^{+\infty}f_{Y|X}(y|x)f_X(x)\mathrm{d}x\\&=\frac{1}{2\pi\sigma r}\int_{-\infty}^{+\infty}\mathrm{e}^{-(\sigma^2+r^2)\left[\left(x-\frac{m\sigma^2+yr^2}{\sigma^2+r^2}\right)^2-\left(\frac{m\sigma^2+y^2r^2}{\sigma^2+r^2}\right)\right]/(2\sigma^2r^2)}\mathrm{d}x\\&=\frac{1}{\sqrt{2\pi(\sigma^2+r^2)}}\mathrm{e}^{-(\sigma^2+r^2)[(m^2\sigma^2+y^2r^2)(\sigma^2+r^2)-(m\sigma^2+yr^2)^2]/[2\sigma^2r^2(\sigma^2+r^2)^2]}\\&=\frac{1}{\sqrt{2\pi(\sigma^2+r^2)}}\mathrm{e}^{-(y-m)^2/[2(\sigma^2+r^2)]}.\end{aligned}$$

例 8 设 (X,Y) 的概率密度为

$$f(x,y)=\begin{cases}x+y, & 0<x<1,0<y<1,\\ 0, & \text{其它},\end{cases}$$

求：在 $0<X<1/n$ 的条件下，Y 的分布函数和概率密度.

解
$$\begin{aligned}&F(y|0<X<1/n)\\&=P\{Y\leqslant y|0<X<1/n\}\\&=P\{0<X<1/n,Y\leqslant y\}/P\{0<X<1/n\}.\end{aligned}$$

当 $y<0$ 时， $F(y<0|0<X<1/n)=0$；

当 $0\leqslant y<1$ 时，

$$P\left\{0<X<\frac{1}{n},Y\leqslant y\right\}=\int_0^{1/n}\mathrm{d}x\int_0^y(x+y)\mathrm{d}y=\frac{y(1+ny)}{2n^2},$$

$$P\left\{0<X<\frac{1}{n}\right\}=\int_0^{1/n}\mathrm{d}x\int_0^1(x+y)\mathrm{d}y=\frac{n+1}{2n^2},$$

所以
$$F\left(y\,\middle|\,0<X<\frac{1}{n}\right)=\begin{cases}0, & y<0,\\ y(1+ny)/(1+n), & 0\leqslant y<1,\\ 1, & y>1,\end{cases}$$

$$f\left(y\,\middle|\,0<X<\frac{1}{n}\right)=\begin{cases}(1+2ny)/(1+n), & 0\leqslant y<1,\\ 0, & \text{其它},\end{cases}$$

例 9 设随机变量 X,Y 都是连续型的，且

$$f_X(x)=\begin{cases}4x\mathrm{e}^{-2x}, & x>0,\\ 0, & x\leqslant 0,\end{cases}$$

$$f_{Y|X}(y|x)=\begin{cases}1/x, & 0<y<x,\\ 0, & \text{其它},\end{cases}$$

求：(1) $f(x,y)$； (2) $f_Y(y)$.

解 因为 $$f(x,y)=f_X(x)f_{Y|X}(y|x),$$

所以
$$f(x,y)=\begin{cases}4\mathrm{e}^{-2x}, & 0<y<x,\\ 0, & \text{其它},\end{cases}$$

$$f_Y(y)=\begin{cases}\displaystyle\int_y^{+\infty}4\mathrm{e}^{-2x}\mathrm{d}x=2\mathrm{e}^{-2y}, & y>0,\\ 0, & y\leqslant 0.\end{cases}$$

例10 设随机变量 X 在 $(0,1)$ 上随机地取值，当 X 取到 x 时，Y 在 $(x,1)$ 上随机地取值，求：

(1) $f(x,y)$； (2) $f_Y(y)$.

解 因为 $X\sim U(0,1)$，在 $X=x$ 的条件下，$Y\sim U(x,1)$，所以

$$f_X(x)=\begin{cases}1, & 0<x<1,\\ 0, & \text{其它},\end{cases}$$

$$f_{Y|X}(y|x)=\begin{cases}1/(1-x), & x<y<1,\\ 0, & \text{其它}.\end{cases}$$

(1) 由 $f(x,y)=f_X(x)f_{Y|X}(y|x)$，得

$$f(x,y)=\begin{cases}1/(1-x), & 0<x<y<1,\\ 0, & \text{其它}.\end{cases}$$

(2) $f_Y(y)=\int_0^y \frac{1}{1-x}\mathrm{d}x=-\ln(1-y)$，故

$$f_Y(y)=\begin{cases}-\ln(1-y), & 0<y<1,\\ 0, & \text{其它}.\end{cases}$$

例 11 设二维随机变量 (X,Y) 的分布函数为

$$F(x,y)=A\left(B+\arctan\frac{x}{2}\right)\left(C+\arctan\frac{y}{3}\right),$$
$$-\infty<x,y<+\infty,$$

求：(1) A,B,C 的值； (2) $f(x,y)$； (3) $f_X(x),f_Y(y)$.

解 因为 $A\neq 0$，所以由 x,y 的任意性，有

$$F(0,-\infty)=A\left(B+\arctan\frac{0}{2}\right)\left(C-\frac{\pi}{2}\right)=0\Longrightarrow C=\frac{\pi}{2},$$

$$F(-\infty,0)=A\left(B-\frac{\pi}{2}\right)\left(C+\arctan\frac{0}{3}\right)=0\Longrightarrow B=\frac{\pi}{2},$$

$$F(+\infty,-\infty)=A\left(\frac{\pi}{2}+\frac{\pi}{2}\right)\left(\frac{\pi}{2}+\frac{\pi}{2}\right)=1\Longrightarrow A=\frac{1}{\pi^2},$$

于是 $$F(x,y)=\frac{1}{\pi^2}\left(\frac{\pi}{2}+\arctan\frac{x}{2}\right)\left(\frac{\pi}{2}+\arctan\frac{y}{3}\right).$$

由 $$f(x,y)=\frac{\partial^2 F(x,y)}{\partial x\partial y},$$

得 $$f(x,y)=\frac{6}{\pi^2(4+x^2)(9+y^2)},\quad -\infty<x,y<+\infty.$$

$$f_X(x)=\int_{-\infty}^{+\infty}f(x,y)\mathrm{d}y=\frac{2}{\pi(4+x^2)},\quad -\infty<x<+\infty,$$

$$f_Y(y)=\int_{-\infty}^{+\infty}f(x,y)\mathrm{d}x=\frac{3}{\pi(9+y^2)},\quad -\infty<y<+\infty.$$

例 12 设随机变量 X,Y 的概率密度分别为 $f_X(x),f_Y(y)$，且

$$f(x,y)=f_X(x)f_Y(y)+h(x,y),\quad -\infty<x,y<+\infty,$$

证明：(1) $h(x,y)\geqslant -f_X(x)f_Y(y)$； (2) $\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}h(x,y)\mathrm{d}x\mathrm{d}y=0$.

证 因为$f(x,y)\geqslant 0$，即$h(x,y)+f_X(x)f_Y(y)\geqslant 0$，所以

$$h(x,y)\geqslant -f_X(x)f_Y(y).$$

又因为
$$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}f(x,y)\mathrm{d}x\mathrm{d}y=1,$$

即
$$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}[f_X(x)f_Y(y)+h(x,y)]\mathrm{d}x\mathrm{d}y=1,$$

所以
$$\begin{aligned}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}h(x,y)\mathrm{d}x\mathrm{d}y&=1-\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}f_X(x)f_Y(y)\mathrm{d}x\mathrm{d}y\\&=1-\int_{-\infty}^{+\infty}f_X(x)\mathrm{d}x\int_{-\infty}^{+\infty}f_Y(y)\mathrm{d}y\\&=1-1=0.\end{aligned}$$

例 13 设

$$f(x,y)=\begin{cases}1, & 0\leqslant x\leqslant 2,\max(0,x-1)\leqslant y\leqslant\min(1,x),\\0, & \text{其它},\end{cases}$$

求：$f_X(x)$和$f_Y(y)$.

解 $\max(0,x-1)=\begin{cases}0, & x<1,\\x-1, & 1\leqslant x,\end{cases}$ $\min(1,x)=\begin{cases}x, & x<1,\\1, & x\geqslant 1,\end{cases}$

所以，$f(x,y)$有意义的区域(见图 3.11)可分为

$$\{0\leqslant x\leqslant 1,0\leqslant y\leqslant x\},$$
$$\{1\leqslant x\leqslant 2,x-1\leqslant y\leqslant 1\},$$

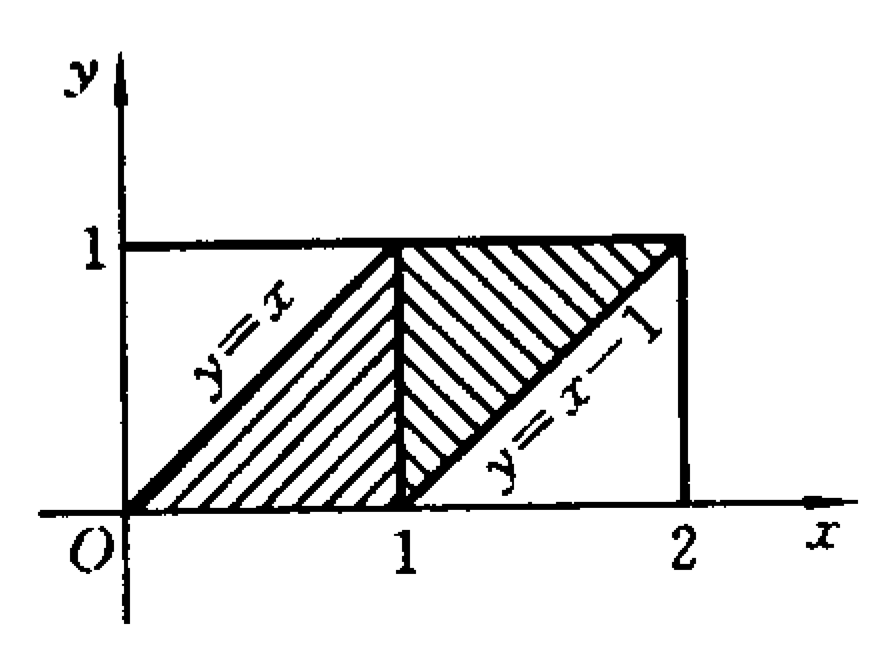


图 3.11

即

$$f(x,y)=\begin{cases}1, & 0\leqslant x<1,0\leqslant y\leqslant x,\\1, & 1\leqslant x\leqslant 2,x-1\leqslant y\leqslant 1,\\0, & \text{其它},\end{cases}$$

所以
$$f_X(x)=\begin{cases}\int_0^1\mathrm{d}y=x, & 0\leqslant x<1,\\\int_{x-1}^1\mathrm{d}y=2-x, & 1\leqslant x\leqslant 2,\\0, & \text{其它},\end{cases}$$

$$f_Y(y)=\begin{cases}\int_y^{y+1}\mathrm{d}x=1, & 0\leqslant y\leqslant 1,\\0, & \text{其它}.\end{cases}$$

第三节　独立性及其应用

主要内容

1. 随机变量的相互独立

设(X,Y)是二维随机变量，如果对于任意的x,y，有$P\{X\leqslant x,Y\leqslant y\}=P\{X\leqslant x\}P\{Y\leqslant y\}$，则称随机变量$X,Y$相互独立.

2. 相互独立的等价命题

若已知$F(x,y)$与$F_X(x)$和$F_Y(y)$，则

$$X,Y\text{ 相互独立}\Longleftrightarrow F(x,y)=F_X(x)F_Y(y).$$

若(X,Y)是连续型随机变量，则

$$X,Y\text{ 相互独立}\Longleftrightarrow f(x,y)=f_X(x)f_Y(y).$$

若(X,Y)是离散型随机变量，则

$$X,Y\text{ 相互独立}\Longleftrightarrow P\{X=x_i,Y=y_j\}=P\{X=x_i\}P\{Y=y_j\}$$

或

$$p_{ij}=p_{i\cdot}\,p_{\cdot j}.$$

对于$(X,Y)\sim N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$，有

$$X,Y\text{ 相互独立}\Longleftrightarrow \rho=0.$$

疑难解析

1. 两个随机变量的相互独立与两个随机事件的相互独立是否相同？为什么？

答　两个随机事件的相互独立是指同一随机试验的同一样本空间上的两个事件的关系，其中一个事件的发生与另一个事件的发生无关，存在$P(AB)=P(A)P(B)$.

两个随机变量的相互独立是指组成二维随机变量(X,Y)的两个向量的关系(但它们也是同一随机试验的同一样本空间上的)，其中一个随机变量的取值与另一个随机变量的取值无关，存在$P\{X\leqslant x,Y\leqslant y\}=P\{X\leqslant x\}P\{Y\leqslant y\}$.

随机变量X和Y是事件的集合，当把$\{X\leqslant x\}$和$\{Y\leqslant y\}$看作两个事件时，两个随机变量相互独立与两个随机事件相互独立是一致的，只是因为X与Y所含事件数多一些，实际上要求也高一些.

特别要注意的是，事件组$\{X_1,X_2,\cdots,X_n\}$的相互独立性是指对$\mathbf{R}$中的任意集合$A_1,A_2,\cdots,A_n$，事件组$\{X_i\in A_i\}$相互独立，而$(X_1,X_2,\cdots,X_n)$的两两独立是指事件组$\{X\in A_i\}$两两独立，与第一章中事件的相互独立与两两独立不相同是一致的.

2. 参数可加性与随机变量独立性有什么样的关系？

答 如果相互独立的两个随机变量X和Y服从的参数分布是相同的，则它们的和$X+Y$也服从同一参数分布，且参数具有可加性.

如$X\sim N(\mu_1,\sigma_1^2)$，$Y\sim N(\mu_2,\sigma_2^2)$，则

$$X+Y\sim N(\mu_1+\mu_2,\sigma_1^2+\sigma_2^2);$$

如
$$X\sim\pi(\lambda_1),\quad Y\sim\pi(\lambda_2),$$
则
$$X+Y\sim\pi(\lambda_1+\lambda_2);$$
如
$$X\sim\chi^2(n_1),\quad Y\sim\chi^2(n_2),$$
则
$$X+Y\sim\chi^2(n_1+n_2).$$
其它如二项分布、泊松分布等都具有此性质.

参数可加性只对相互独立的随机变量适用，它为讨论问题和进行计算带来很大方便. 我们要学会利用这一性质.

方法、技巧与典型例题分析

判别两个随机变量的独立性的方法大体有三种：

(1) 由定义和它的等价命题来判别，这时必先求得边缘分布与联合分布，然后进行验证.

(2) 利用微积分中的性质，若 $f(x,y)\in G$，且有

$$f(x,y)=g(x)h(y),\quad x,y\in G,$$

其中 $g(x)$，$h(y)$ 是 x，y 的非负可积函数，则组成 (X,Y) 的随机变量 X 与 Y 相互独立.

(3) 利用对称性和经验，确定随机变量的相互独立.

例 1 设随机变量 X 以概率 1 取值为零，而 Y 是任意的随机变量，证明 X 与 Y 相互独立.

解 因为 X 的分布函数为

$$F(x)=\begin{cases}0, & x<0,\\ 1, & x\geqslant 0,\end{cases}$$

设 Y 的分布函数为 $F_Y(y)$，(X,Y) 的分布函数为 $F(x,y)$，则：

当 $x<0$ 时，对任意 y，有

$$\begin{aligned}F(x,y)&=P\{X\leqslant x,Y\leqslant y\}=P\{X\leqslant x\cap Y\leqslant y\}\\&=P\{\varnothing\cap Y\leqslant y\}=P\{\varnothing\}=0=F_X(x)F_Y(y);\end{aligned}$$

当 $x\geqslant 0$ 时，对任意 y，有

$$\begin{aligned}F(x,y)&=P\{X\leqslant x,Y\leqslant y\}=P\{X\leqslant x\cap Y\leqslant y\}\\&=P\{Y\leqslant y\}=F_Y(y)=F_X(x)F_Y(y).\end{aligned}$$

依定义，由 $F(x,y)=F_X(x)F_Y(y)$ 知，X 与 Y 相互独立.

例 2 设 $X\sim B(n_1,p)$，$Y\sim B(n_2,p)$，证明：X，Y 相互独立，则

$$X+Y\sim B(n_1+n_2,p).$$

证 因为
$$P\{X=i\}=C_{n_1}^{i}p^{i}q^{n_1-i}\quad(i=0,1,\cdots,n_1),$$
$$P\{Y=j\}=C_{n_2}^{j}p^{i}q^{n_2-j}\quad(j=0,1,\cdots,n_2),$$

所以

$$\begin{aligned}P\{X+Y=k\}&=\sum_{i=0}^{k}P\{X=i,Y=k-i\}\\&=\sum_{i=0}^{k}P\{X=i\}P\{Y=k-i\}\\&=\sum_{i=0}^{k}(C_{n_1}^{i}p^{i}q^{n_1-i})(C_{n_2}^{k-i}p^{k-i}q^{n_2-k+i})\end{aligned}$$

$$=\sum_{i=0}^{k}C_{n_1}^{i}C_{n_2}^{k-i}p^k q^{n_1+n_2-k}=\left(\sum_{i=0}^{k}C_{n_1}^{i}C_{n_2}^{k-i}\right)p^k q^{n_1+n_2-k}$$

$$=C_{n_1+n_2}^{k}p^k q^{n_1+n_2-k}\quad(k=0,1,\cdots,n_1+n_2),$$

于是
$$X+Y\sim B(n_1+n_2,p).$$

例 3 设 X 与 Y 相互独立，$X\sim\pi(\lambda_1)$，$Y\sim\pi(\lambda_2)$，证明：

$$X+Y\sim\pi(\lambda_1+\lambda_2).$$

证 因为 $P\{X=i\}=\lambda_1^i e^{-\lambda_1}/i!\ (i=0,1,\cdots)$，

$$P\{Y=j\}=\lambda_2^j e^{-\lambda_2}/j!\ (j=0,1,\cdots),$$

所以 $P\{X+Y=k\}$

$$=\sum_{j=0}^{k}P\{X=i,Y=k-i\}=\sum_{i=0}^{k}P\{X=i\}P\{Y=k-i\}$$

$$=\sum_{i=0}^{k}\frac{\lambda_1^i\lambda_2^{k-i}}{i!\ (k-i)!}e^{-\lambda_1-\lambda_2}=\frac{e^{-(\lambda_1+\lambda_2)}}{k!}\sum_{i=0}^{k}\frac{k!}{i!\ (k-i)!}\lambda_1^i\lambda_2^{k-i}$$

$$=\frac{(\lambda_1+\lambda_2)^k}{k!}e^{-(\lambda_1+\lambda_2)},\quad k=0,1,2,\cdots,$$

于是
$$X+Y\sim\pi(\lambda_1+\lambda_2).$$

例 4 设随机变量 (X,Y) 的联合分布律为

Y \ X	x_1	x_2	x_3
y_1	a	1/9	c
y_2	1/9	b	1/3

若 X,Y 相互独立，求 a,b,c 的值.

解 因为 X,Y 的边缘分布律分别为

X	x_1	x_2	x_3
p_k	$a+1/9$	$b+1/9$	$c+1/3$

Y	y_1	y_2
p_k	$a+c+1/9$	$b+4/9$

$$p_{22}=(b+1/9)(b+4/9)=b\Longrightarrow b=2/9,$$

$$p_{12}=(a+1/9)(b+4/9)=1/9\Longrightarrow a=1/18,$$

$$\sum_i\sum_j p_{ij}=a+b+c+1/9+1/9+1/3=1\Longrightarrow c=1/6.$$

经验证，$a=1/18$，$b=2/9$，$c=1/6$ 确为本题的解.

例5 设随机变量(X,Y)的联合分布律为

X \ Y	0	1	2
0	0.06	0.15	α
1	β	0.35	0.21

问：当α,β为何值时，X与Y相互独立？

解 X与Y的边缘分布律分别为

X	0	1
p_k	$0.21+\alpha$	$0.56+\beta$

Y	0	1	2
p_k	$0.06+\beta$	0.5	$0.21+\alpha$

若$p_{ij}=p_{i\cdot}\,p_{\cdot j}$，则$X$与$Y$相互独立. 又

$$0.15=p_{01}=p_{0\cdot}\,p_{\cdot 1}=0.5(\alpha+0.21)\Longrightarrow\alpha=0.09,$$

$$0.35=p_{11}=p_{1\cdot}\,p_{\cdot 1}=0.5(\beta+0.56)\Longrightarrow\beta=0.14.$$

经验证，$\alpha=0.09,\beta=0.14$确为本题的解，此时X与Y相互独立.

例6 设二维随机变量(X,Y)的联合密度为

$$f(x,y)=\begin{cases}(1+xy)/4, & |x|<1,|y|<1,\\ 0, & 其它,\end{cases}$$

证明：X与Y不相互独立，但X^2与Y^2相互独立.

证 当$|x|<1$时，$f_X(x)=\int_{-1}^{1}\frac{1+xy}{4}\mathrm{d}y=\frac{1}{2}$；当$|x|\geqslant 1$时，$f_X(x)=0$. 故

$$f_X(x)=\begin{cases}1/2, & |x|<1,\\ 0, & 其它.\end{cases}$$

类似地，
$$f_Y(y)=\begin{cases}1/2, & |y|<1,\\ 0, & 其它.\end{cases}$$

显然，当$0<|x|<1,0<|y|<1$时，$f(x,y)\neq f_X(x)\cdot f_Y(y)$，所以$X$与$Y$不相互独立.

设X^2的分布函数为$F_1(x)$，在$0\leqslant x\leqslant 1$内，

$$F_1(x)=P\{X^2\leqslant x\}=\int_{-\sqrt{x}}^{\sqrt{x}}\frac{1}{2}\mathrm{d}x=\sqrt{x},$$

即

$$F_1(x)=\begin{cases}0, & x<0,\\ \sqrt{x}, & 0\leqslant x<1,\\ 1, & x\geqslant 1,\end{cases}$$

$$F_2(y)=\begin{cases}0, & y<0,\\ \sqrt{y}, & 0\leqslant y<1,\\ 1, & y\geqslant 1.\end{cases}$$

设(X^2,Y^2)的分布函数为$F_3(x,y)$，则：
当$x<0$或$y<0$时，$F_3(x,y)=0$；
当$0\leqslant x<1,y\geqslant 1$时，

$$F_3(x,y)=P\{X^2\leqslant x,Y^2\leqslant y\}=P\{X^2\leqslant x\}=\sqrt{x};$$

当$0\leqslant y<1,x\geqslant 1$时，$F_3(x,y)=\sqrt{y}$；
当$0\leqslant x<1,0\leqslant y<1$时，

$$F_3(x,y)=\int_{-\sqrt{x}}^{\sqrt{x}}\mathrm{d}x\int_{-\sqrt{y}}^{\sqrt{y}}\frac{1+xy}{4}\mathrm{d}y=\sqrt{xy};$$

当$x\geqslant 1,y\geqslant 1$时，$F_3(x,y)=1$.

所以

$$F_3(x,y)=\begin{cases}0, & x<0\text{ 或 }y<0,\\ \sqrt{x}, & 0\leqslant x<1,y\geqslant 1,\\ \sqrt{y}, & 0\leqslant y<1,x\geqslant 1,\\ \sqrt{xy}, & 0\leqslant x<1,0\leqslant y<1,\\ 1, & x\geqslant 1,y\geqslant 1.\end{cases}$$

经验证，$F_3(x,y)=F_1(x)F_2(y)$对所有x,y成立，所以X^2与Y^2相互独立.

例 7 设随机变量(X,Y)的概率密度为

$$f(x,y)=A\mathrm{e}^{-ax^2+bxy-cy^2},\quad -\infty<x,y<+\infty,$$

问：a,b,c满足什么条件时，X与Y相互独立？

解 X与Y的边缘概率密度为

$$f_X(x)=\int_{-\infty}^{+\infty}A\mathrm{e}^{-ax^2+bxy-cy^2}\mathrm{d}y=A\mathrm{e}^{-ax^2}\mathrm{e}^{[bx/(2\sqrt{c})]^2}\int_{-\infty}^{+\infty}\mathrm{e}^{-t^2}\mathrm{d}t$$

$$=A\sqrt{\pi/c}\mathrm{e}^{-[a-b^2/(4c)]x^2},\quad -\infty<x<+\infty$$

（令 $t=\sqrt{c}\,y-bx/(2\sqrt{c})$），

类似地，　　$f_Y(y)=A\sqrt{\pi/a}\mathrm{e}^{-[c-b^2/(4a)]y^2}$，$-\infty<y<+\infty$.

X,Y 相互独立，应该有 $f(x,y)=f_X(x)f_Y(y)$，即

$$A\mathrm{e}^{-ax^2+bxy-cy^2}=A^2\pi/\sqrt{ac}\cdot \mathrm{e}^{-[a-b^2/(4c)]x^2-[c-b^2/(4a)]y^2}.$$

比较等式两边系数知，应该满足条件.

$$b=0,\quad \sqrt{ac}=A\pi.$$

例8　设随机变量(X,Y)的两个分量X和Y相互独立，且服从同一分布，试证：$P\{X\leqslant Y\}=1/2$.

证　因为X,Y相互独立，所以$f(x,y)=f_X(x)f_Y(y)$. 于是

$$\begin{aligned}P\{X\leqslant Y\}&=\iint\limits_{x\leqslant y}f(x,y)\mathrm{d}x\mathrm{d}y=\iint\limits_{x\leqslant y}f_X(x)f_Y(y)\mathrm{d}x\mathrm{d}y\\&=\int_{-\infty}^{+\infty}\left[f_Y(y)\int_{-\infty}^{y}f_X(x)\mathrm{d}x\right]\mathrm{d}y\\&=\int_{-\infty}^{+\infty}[f_Y(y)F_Y(y)]\mathrm{d}y\\&=\int_{-\infty}^{+\infty}F_Y(y)\mathrm{d}F_Y(y)=\frac{1}{2}F^2(y)\Big|_{-\infty}^{+\infty}=\frac{1}{2}.\end{aligned}$$

也可以利用对称性来证. 因为X,Y相互独立且同分布，所以

$$P\{X\leqslant Y\}=P\{Y\leqslant X\},$$

而$P\{X\leqslant Y\}+P\{X\geqslant Y\}=1$，故$P\{X\leqslant Y\}=1/2$.

例9　设随机变量X与Y相互独立，且

$$P\{X=1\}=P\{Y=1\}=p>0,$$
$$P\{X=0\}=P\{Y=0\}=1-p>0,$$

定义随机变量

$$Z=\begin{cases}1, & X+Y\text{ 为偶数},\\ 0, & X-Y\text{ 为奇数},\end{cases}$$

问：p为何值时，X与Z相互独立？

解　先写出(X,Y,Z)的联合分布律，再写出(X,Z)的联合分

布律(用求边缘分布律的方法).

当 $z=1$ 时，

Y \ X	0	1
0	$(1-p)^2$	0
1	0	p^2

当 $z=0$ 时，

Y \ X	0	1
0	0	$p(1-p)$
1	$p(1-p)$	0

所以，(X,Z)的联合分布律为

X \ Z	0	1
0	$p(1-p)$	$(1-p)^2$
1	$p(1-p)$	p^2

对 $X=i,Z=j$，且 $i,j=0,1$，若 X,Z 相互独立，应有

$$p_{ij}=p_{i\cdot}\,p_{\cdot j}.$$

由 $$P\{X=1,Z=1\}=p^2=P\{X=1\}P\{Z=1\}=p[p^2+(1-p)^2]$$

解得 $p=1/2$. 经验证，对(X,Z)的一切(i,j)，均满足 $p_{ij}=p_{i\cdot}\,p_{\cdot j}$，所以，当 $p=1/2$ 时，X 与 Z 相互独立.

例 10 设 X 与 Y 相互独立，$X\sim U[a,b]$，$Y\sim e(\lambda)$，求：

(1) $f(x,y)$； (2) 概率 $P\{Y\leqslant X\}$.

解 $$f_X(x)=\begin{cases}1/(b-a), & a\leqslant x\leqslant b,\\ 0, & \text{其它},\end{cases}$$

$$f_Y(y)=\begin{cases}\lambda e^{-\lambda y}, & y\geqslant 0,\\ 0, & \text{其它},\end{cases}$$

所以 $$f(x,y)=\begin{cases}\dfrac{\lambda e^{-\lambda y}}{b-a}, & a\leqslant x\leqslant b, y\geqslant 0,\\ 0, & \text{其它},\end{cases}$$

$$P\{Y\leqslant X\}=\iint\limits_{y\leqslant x}\frac{\lambda e^{-\lambda y}}{b-a}dxdy=\frac{\lambda}{b-a}\int_a^b dx\int_0^x e^{-\lambda y}dy$$

$$=\frac{1}{b-a}\int_a^b(1-e^{-\lambda x})dx=1+\frac{1}{(b-a)\lambda}(e^{-\lambda b}-e^{-\lambda a}).$$

例 11 任意取两个正的真分式，求其积不大于 2/9 且其和不大1 的概率.

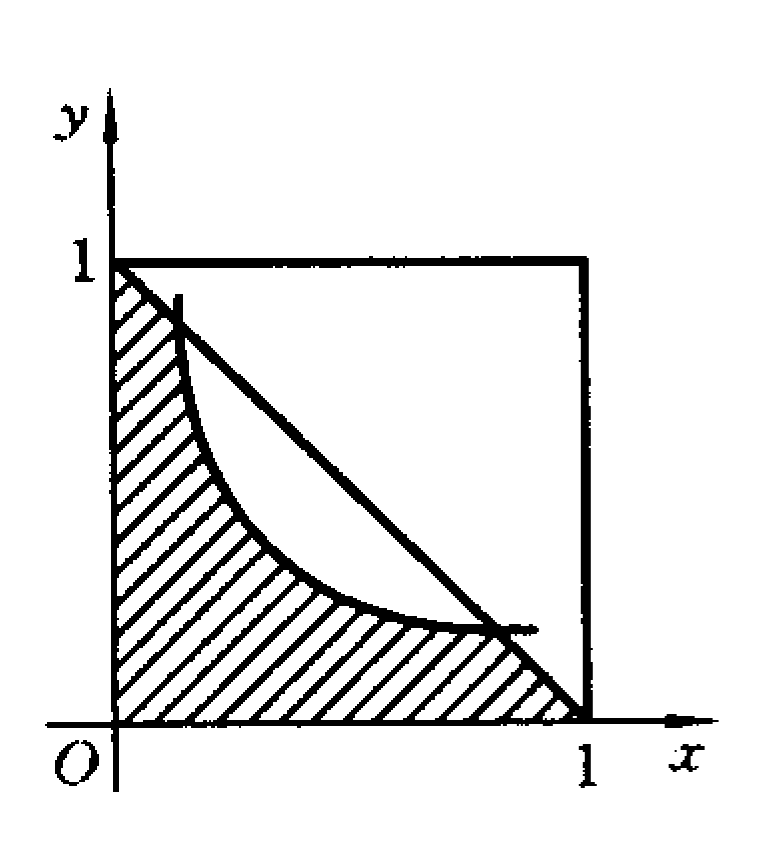


图 3.12

解 以随机变量 X 和 Y 记所取的两个真分式，则可取值 $0<x<1, 0<y<1$（见图 3.12），所求概率为

$$P\{(x,y)\in\{xy\leqslant 2/9, x+y\leqslant 1\}\},$$

且 X,Y 相互独立. 因为

$$f(x,y)=\begin{cases}1, & 0<x<1, 0<y<1,\\ 0, & \text{其它},\end{cases}$$

所以 $$P\{(x,y)\in\{xy\leqslant 2/9, x+y\leqslant 1\}\}$$
$$=\frac{1}{2}-\int_{1/3}^{2/3}\mathrm{d}x\int_{2/9x}^{1-x}\mathrm{d}y=\frac{1}{2}-\int_{1/3}^{2/3}\frac{1-x-2}{9x}\mathrm{d}x$$
$$=\frac{1}{3}+\frac{2}{9}\ln 2=0.4873.$$

例 12 设 X 与 Y 是相互独立的随机变量，且服从同一分布 $U(-a,a]$ $(a>0)$，求方程 $t^2+Xt+Y=0$ 有实根的概率，并求 $a\to 0$ 和 $a\to\infty$ 时，此概率的极限值.

解 $$f(x,y)=\begin{cases}1/(4a^2), & |x|\leqslant a, |y|\leqslant a,\\ 0, & \text{其它}.\end{cases}$$

要方程有实根，应该有 $X^2-4Y\geqslant 0$，要求出

$$P\{X^2\geqslant 4Y\}=\iint\limits_{x^2\geqslant 4y} f(x,y)\mathrm{d}x\mathrm{d}y.$$

当 $0\leqslant a\leqslant 4$ 时，由图 3.13(a)知

$$P\{X^2\geqslant 4Y\}=2\int_0^a\mathrm{d}x\int_{-a}^{x^2/4}\frac{1}{4a^2}\mathrm{d}y=\frac{a}{24}+\frac{1}{2};$$

当 $a>4$ 时，由图 3.13(b)知

$$P\{X^2\geqslant 4Y\}=1-2\int_0^a\mathrm{d}y\int_0^{2\sqrt{y}}\frac{1}{4a^2}\mathrm{d}x=1-\frac{2}{3\sqrt{a}}.$$

于是 $$\lim_{a\to 0}P\{X^2\geqslant 4Y\}=\lim_{a\to 0}\frac{a}{24}+\frac{1}{2}=\frac{1}{2},$$

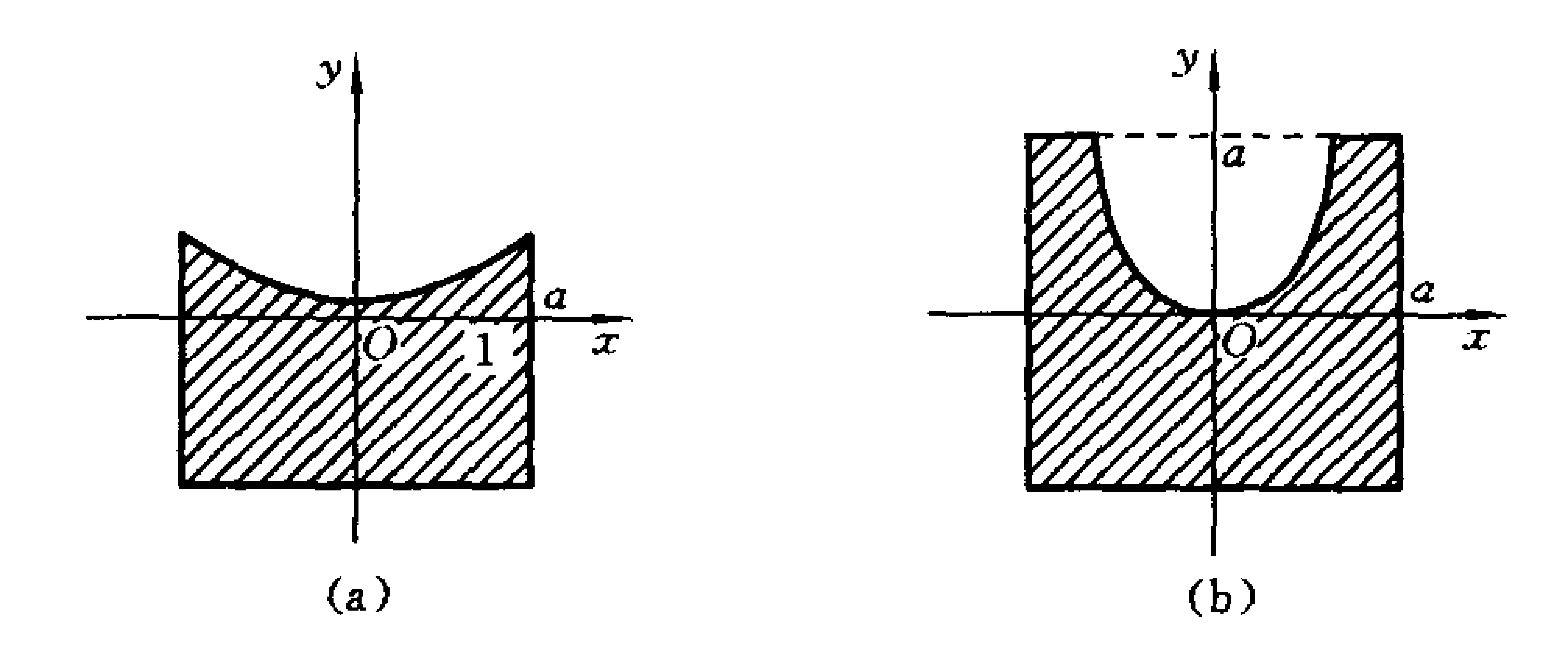


图 3.13

$$\lim_{a\to\infty}P\{X^2\geqslant 4Y\}=\lim_{a\to\infty}\left(1-\frac{2}{3\sqrt{a}}\right)=1.$$

第四节　两个随机变量的函数的分布

主要内容

两个随机变量的函数的分布，形式较多，只要求掌握以下几种即可.

1. 两个随机变量的和 $Z=X+Y$ 的分布

设已知随机变量 X 和 Y 的概率分布为 $P\{X=x_i\}$ 和 $P\{Y=y_j\}$，(X,Y) 的联合分布为 $P\{X=x_i,Y=y_j\}$，则

$$P\{Z=k\}=P\{X+Y=k\}=\sum_{i=1}^{k}P\{X=i,Y=k-i\}.$$

当 X,Y 相互独立时，

$$P\{Z=k\}=\sum_{i=1}^{k}P\{X=i\}P\{Y=k-i\}.$$

设 (X,Y) 是连续型随机变量，(X,Y) 的概率密度为 $f(x,y)$，则对任意实数 z，有

$$F_Z(z)=\int_{-\infty}^{z}\left[\int_{-\infty}^{+\infty}F(u-y,y)\mathrm{d}y\right]\mathrm{d}u,$$

$$f_Z(z)=\int_{-\infty}^{+\infty}f(z-y,y)\mathrm{d}y=\int_{-\infty}^{+\infty}f(x,z-x)\mathrm{d}x.$$

当 X 与 Y 相互独立时，有卷积公式

$$f_Z(z)=f_X*f_Y=\int_{-\infty}^{+\infty}f_X(z-y)f_Y(y)\mathrm{d}y$$

$$=\int_{-\infty}^{+\infty}f_X(x)f_Y(z-x)\mathrm{d}x.$$

2. 两个随机变量的商 $Z=X/Y$ 的分布

设已知 (X,Y) 的概率密度 $f(x,y)$，则 $Z=X/Y$ 的分布函数为

$$F_Z(z)=\int_{-\infty}^{z}\mathrm{d}u\int_{0}^{+\infty}yf(uy,y)\mathrm{d}y-\int_{-\infty}^{z}\mathrm{d}u\int_{-\infty}^{0}yf(uy,y)\mathrm{d}y,$$

Z 的概率密度为

$$f_Z(z)=\int_{-\infty}^{+\infty}|y|f(yz,y)\mathrm{d}y.$$

当 X,Y 相互独立时，上式化为

$$f_Z(z)=\int_{-\infty}^{+\infty}|y|f_X(yz)f_Y(y)\mathrm{d}y.$$

3. 两个随机变量的最大值 $M=\max(X,Y)$ 和最小值 $N=\min(X,Y)$ 的分布

当 X 和 Y 相互独立，且分布函数 $F_X(x)$ 和 $F_Y(y)$ 为已知时，有

$$F_M(z)=F_X(z)F_Y(z),$$

$$F_N(z)=1-[1-F_X(z)][1-F_Y(z)].$$

以上结果可推广到有限个相互独立的随机变量的情形. 若 $X_1,X_2,\cdots,X_n$ 相互独立，分布函数分别为 $F_{x_1}(x_1),F_{x_2}(x_2),\cdots,F_{x_n}(x_n)$，则

$$F_M(z)=F_{X_1}(z)F_{X_2}(z)\cdots F_{X_n}(z),$$

$$F_N(z)=1-[1-F_{X_1}(z)]\cdots[1-F_{X_n}(z)].$$

当 $X_1,X_2,\cdots,X_n$ 相互独立且同分布时，有

$$F_M(z)=[F(z)]^n,\quad F_N(z)=1-[1-F(z)]^n.$$

疑难解析

1. 两个相互独立的服从正态分布的随机变量X_1与X_2之和仍是正态分布的随机变量,那么它们的线性组合呢?

答 由上节知,有限个正态分布的随机变量之和是正态随机变量,而且具有参数可加性,则

$$aX_1 \sim N(a\mu_1, a\sigma_1^2), \quad bX_2 \sim N(b\mu_2, b\sigma_2^2),$$

所以

$$aX_1 + bX_2 \sim N(a\mu_1 + b\mu_2, a\sigma_1^2 + b\sigma_2^2),$$

即两个正态随机变量的线性组合仍是正态随机变量,且它们的参数是相应参数的线性组合. 严格的证明可以用分布函数法作出(a, b 为正整数).

2. 用卷积公式计算$Z=X+Y$的密度函数时要注意些什么?

答 当X与Y相互独立,且密度函数分别为$f_X(x)$和$f_Y(y)$时,有卷积公式

$$f_Z(z) = \int_{-\infty}^{+\infty} f_X(x) f_Y(z-x)\,\mathrm{d}x = \int_{-\infty}^{+\infty} f_X(z-y) f_Y(y)\,\mathrm{d}y.$$

同以前的积分一样,我们要进行仔细的讨论:即当z在不同区间取值,被积函数的形式是否相同,积分是否要分段进行,特别要排除被积函数为零的区间.

当X,Y相互独立,且X,Y的密度函数分别为$f_X(x)$和$f_Y(y)$时,$Z=aX+bY$的密度函数公式为

$$f_Z(z) = \int_{-\infty}^{+\infty} f_X(x) \frac{1}{|b|} f_Y\left(\frac{z-ax}{b}\right)\mathrm{d}x$$

或

$$f_Z(z) = \int_{-\infty}^{+\infty} \frac{1}{|a|} f_X\left(\frac{z-by}{a}\right) f_Y(y)\,\mathrm{d}y.$$

方法、技巧与典型例题分析

求两个随机变量函数的分布通常有两种方法：

一种方法是分布函数法，即设 $Z=g(X,Y)$，先求出 $F_Z(z)$，再求 $f_Z(z)$. 教材中已给出 $Z=X+Y$，$Z=X/Y$，$Z=\max(X,Y)$ 与 $Z=\min(X,Y)$ 的有关公式. 但要注意，有些公式对 X,Y 是有要求的，如卷积公式、$Z=\max(X,Y)$ 与 $Z=\min(X,Y)$ 公式仅当 X 与 Y 相互独立时才适用. 同时，在求积分时，则要注意分区域积分和正确配置积分限.

另一种方法是引入随机变量函数组（即坐标变换）法，其一般步骤是：

(1) 建立随机变量函数组 $\begin{cases} V=g(X,Y), \\ U=X \text{ 或 } Y \text{ 或 } h(X,Y), \end{cases}$ 并写出逆变换式；

(2) 求出变换的雅可比行列式 $J=\partial(xy)/\partial(u,v)$；

(3) 写出新随机变量组 (U,V) 的联合密度函数

$$f_{UV}(u,v)=f[x(u,v),y(u,v)]\cdot|J|;$$

(4) 求出边缘概率密度 $f_V(u)$.

两种方法各有优缺点. 分布函数法便于掌握，但计算二重积分时要讨论积分区域，比较麻烦. 引入随机变量函数组法使变换后的密度函数的积分区域易于确定，积分容易计算，但要引入另一随机变量.

例1 设 $X\sim N(0,1)$，$Y\sim N(0,1)$，且 X 与 Y 相互独立，求 $Z=X+Y$ 的分布.

解 因为 X,Y 相互独立，且有

$$f_T(t)=\frac{1}{\sqrt{2\pi}}e^{-t^2/2},\quad -\infty<t<+\infty.$$

利用卷积公式，得

$$f_Z(z)=\int_{-\infty}^{+\infty}\frac{1}{2\pi}e^{-x^2/2}e^{-(z-x)^2/2}dx$$
$$=\frac{1}{2\pi}e^{-z^2/4}\int_{-\infty}^{+\infty}e^{-(x-z/2)^2/2}dx.$$

令 $t=x-z/2$ 得

$$f_Z(z)=e^{-z^2/4}\frac{1}{2\pi}\int_{-\infty}^{+\infty}e^{-t^2}dt=\frac{1}{2\sqrt{\pi}}e^{-z^2/4},-\infty<z<+\infty,$$

所以 $Z\sim N(0,2)$.

例 2 设随机变量 X 和 Y 相互独立，且

$$f_X(x)=\begin{cases}1, & 0\leqslant x\leqslant 1,\\ 0, & \text{其它},\end{cases}\qquad f_Y(y)=\begin{cases}2y, & 0\leqslant y\leqslant 1,\\ 0, & \text{其它},\end{cases}$$

求随机变量 $Z=X+Y$ 的概率密度 $f_Z(z)$.

解 如图 3.14 所示，用分布函数法求解.

$$F_Z(z)=P\{X+Y\leqslant z\}=P\{X+Y\in G:X+Y\leqslant z\}$$
$$=\iint_G f(x,y)dxdy=\iint_G 1\times 2ydxdy,$$

当 $z<0$ 时，　　$F_Z(z)=0$;

当 $0\leqslant z<1$ 时，

$$F_Z(z)=\int_0^z dx\int_0^{z-x}2ydy=\frac{z^3}{3};$$

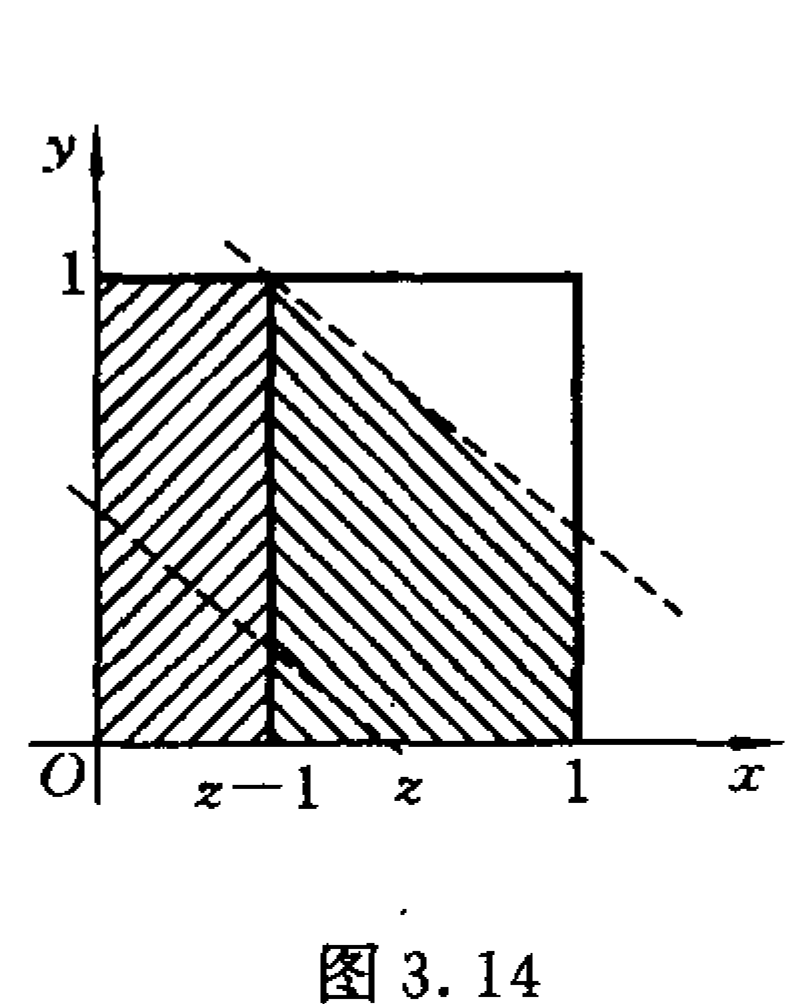


图 3.14

当 $1\leqslant z<2$ 时，

$$F_Z(z)=\int_0^{z-1}dx\int_0^1 2ydy+\int_{z-1}^1 dx\int_0^{z-x}2ydy$$
$$=z^2-z^3/3-1/3;$$

当 $z\geqslant 2$ 时，$F_Z(z)=1$.

由 $F'_Z(z)=f_Z(z)$，得

$$f_Z(z)=\begin{cases}z^2, & 0\leqslant z<1,\\ 2z-z^2, & 1\leqslant z<2,\\ 0, & \text{其它}.\end{cases}$$

例 3 设随机变量 X,Y 相互独立且同分布

$$f_T(t)=\frac{1}{2}\mathrm{e}^{-|t|},\quad -\infty<t<+\infty,$$

求 $Z=X+Y$ 的分布.

解 由卷积公式,当 $z>0$ 时,

$$\begin{aligned}f_Z(z)&=\frac{1}{4}\left[\int_{-\infty}^{0}\mathrm{e}^{x-(z-x)}\mathrm{d}x+\int_{0}^{z}\mathrm{e}^{-x-(z-x)}\mathrm{d}x+\int_{z}^{+\infty}\mathrm{e}^{-x-(x-z)}\mathrm{d}x\right]\\&=\frac{1}{4}\left(\int_{-\infty}^{0}\mathrm{e}^{2x-z}\mathrm{d}x+\int_{0}^{z}\mathrm{e}^{-z}\mathrm{d}x+\int_{z}^{+\infty}\mathrm{e}^{-2x+z}\mathrm{d}x\right)\\&=\frac{1}{4}\left(\frac{\mathrm{e}^{-z}}{2}+z\mathrm{e}^{-z}+\frac{\mathrm{e}^{-z}}{2}\right)=\frac{1}{4}(1+z)\mathrm{e}^{-z}.\end{aligned}$$

由于 $f_Z(-z)=f_Z(z)$,所以

$$f_Z(z)=\frac{1}{4}(1+|z|)\mathrm{e}^{-|z|},\quad -\infty<z<+\infty.$$

例4 设市场上某商品的每周需求量 T 是一个随机变量,密度函数为

$$f_T(t)=\begin{cases}t\mathrm{e}^{-t}, & x>0,\\ 0, & \text{其它}.\end{cases}$$

设各周的需求量是相互独立的,求第 2 周、第 3 周需求量的密度函数.

解 以 $T_i(i=1,2,3)$ 记第 i 周的需求量,T_i 相互独立且同分布,以 X 记 T_1+T_2,则 $x>0$ 时,

$$\begin{aligned}f_X(x)&=\int_{0}^{+\infty}f_{T_1}(t_1)f_{T_2}(x-t_1)\mathrm{d}t_1\\&=\int_{0}^{x}(x-t_1)\mathrm{e}^{-(x-t_1)}\cdot t_1\mathrm{e}^{-t_1}\mathrm{d}t_1=x^3\mathrm{e}^{-x}/6,\end{aligned}$$

所以

$$f_X(x)=\begin{cases}x^3\mathrm{e}^{-x}/6, & x>0,\\ 0, & \text{其它}.\end{cases}$$

以 Y 记 $T_1+T_2+T_3$,则 $y>0$ 时,

$$\begin{aligned}f_Y(y)&=\int_{0}^{+\infty}f_X(x)f_{T_3}(y-x)\mathrm{d}x\\&=\int_{0}^{y}\frac{1}{6}x^3\mathrm{e}^{-x}(y-x)\mathrm{e}^{-(y-x)}\mathrm{d}x=y^5\mathrm{e}^{-y}/120,\end{aligned}$$

所以 $$f_Y(y)=\begin{cases} y^5 e^{-y}/120, & y>0, \\ 0, & y\leqslant 0. \end{cases}$$

例 5 设随机变量 X 与 Y 相互独立，且 $X\sim U[-a,a]$，$Y\sim N(b,\sigma^2)$，求 $Z=X+Y$ 的概率密度.

解 因为

$$f_X(x)=\begin{cases} \dfrac{1}{2a}, & |x|\leqslant a, \\ 0, & \text{其它}, \end{cases} \quad f_Y(y)=\frac{1}{\sqrt{2\pi}b}e^{-(y-b)^2/(2\sigma^2)},$$

所以

$$\begin{aligned} f_Z(z)&=\int_{-\infty}^{+\infty} f_X(x)f_Y(z-x)\mathrm{d}x=\int_{-a}^{a}\frac{1}{2a\sqrt{2\pi}}e^{-(z-x-b)^2/(2\sigma^2)}\mathrm{d}x \\ &=-\frac{1}{2\sqrt{2\pi}a}\int_{(z+a-b)/\sigma}^{(z-a-b)/\sigma} e^{-t^2/2}\mathrm{d}t=\frac{1}{2\sqrt{2\pi}a}\int_{(z-a-b)/\sigma}^{(z+a-b)/\sigma} e^{-t^2/2}\mathrm{d}t \\ &=\frac{1}{2a}\left[\Phi\left(\frac{z+a-b}{\sigma}\right)-\Phi\left(\frac{z-a-b}{\sigma}\right)\right]. \end{aligned}$$

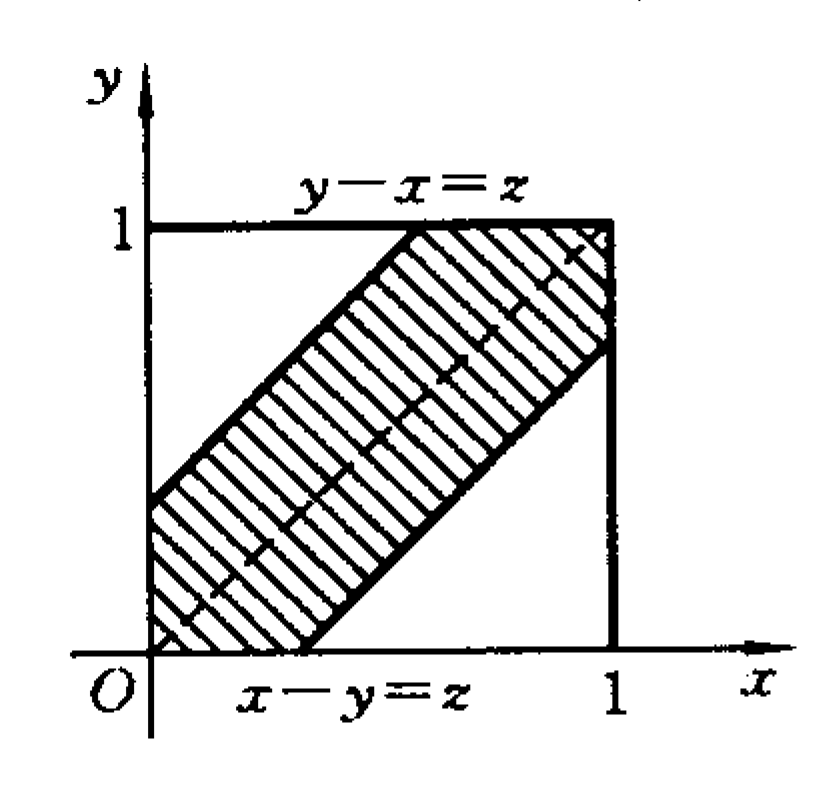


图 3.15

例 6 在区间[0,1]上随机地取得两点 x 和 y（见图 3.15），求这两点间距离的概率密度函数.

解 $(X,Y)\sim U[0,1;0,1]$，联合密度为

$$f(x,y)=\begin{cases} 1, & 0\leqslant x\leqslant 1, 0\leqslant y\leqslant 1, \\ 0, & \text{其它}. \end{cases}$$

以 $Z=|X-Y|$ 记两点 x 与 y 间距离，则

$$F_Z(z)=P\{|X-Y|\leqslant z\}=\iint\limits_{|x-y|\leqslant z} f(x,y)\mathrm{d}x\mathrm{d}y.$$

当 $z<0$ 时， $$F_Z(z)=0;$$

当 $0\leqslant z<1$ 时，

$$F_Z(z)=\iint\limits_D \mathrm{d}x\mathrm{d}y=1-(1-z)^2=2z-z^2;$$

当 $z\geqslant 1$ 时， $$F_Z(z)=1.$$

故
$$F_Z(z)=\begin{cases}0, & z<0,\\ 2z-z^2, & 0\leqslant z<1,\\ 1, & z\geqslant 1,\end{cases}$$
所以，$Z=|X-Y|$的概率密度为
$$f_Z(z)=\begin{cases}2(1-z), & 0\leqslant z<1,\\ 0, & \text{其它}.\end{cases}$$

例7 设随机变量X和Y相互独立，$X\sim N(0,1)$，$Y\sim U(0,1)$，求$Z=X/Y$的概率密度函数.

解 因为 $f_X(x)=\dfrac{1}{\sqrt{2\pi}}e^{-x^2/2}$， $-\infty<x<+\infty$，
$$f_Y(y)=\begin{cases}1, & 0\leqslant y\leqslant 1,\\ 0, & \text{其它},\end{cases}$$
所以 $f_Z(z)=\displaystyle\int_0^1 yf_X(yz)f_Y(y)\mathrm{d}y=\int_0^1\frac{1}{\sqrt{2\pi}}ye^{-(yz)^2/2}\mathrm{d}y.$

由于$f_Z(z)$是z的偶函数，当$z>0$时，可令$x=zy$，于是
$$f_Z(z)=\frac{1}{\sqrt{2\pi}z^2}\int_0^z xe^{-x^2/2}\mathrm{d}x=\frac{1}{\sqrt{2\pi}z^2}(1-e^{-z^2/2}),$$
当$z\leqslant 0$时， $f_Z(z)=\dfrac{1}{\sqrt{2\pi}}\displaystyle\int_0^1 y\mathrm{d}y=\frac{1}{2\sqrt{2\pi}}$，

即
$$f_Z(z)=\begin{cases}\dfrac{1}{\sqrt{2\pi}z^2}(1-e^{-z^2/2}), & z>0,\\ \dfrac{1}{2\sqrt{2\pi}}, & z\leqslant 0.\end{cases}$$

例8 设随机变量X与Y相互独立，且密度函数为
$$f_X(x)=\begin{cases}e^{-x}, & x>0,\\ 0, & x\leqslant 0,\end{cases}\quad f_Y(y)=\begin{cases}2e^{-y}, & y>0,\\ 0, & y\leqslant 0,\end{cases}$$
求随机变量$Z=X/Y$的密度函数.

解 $f_Z(z)=\displaystyle\int_{-\infty}^{+\infty}|y|f_X(yz)f_Y(y)\mathrm{d}y=\int_0^{+\infty}2ye^{-2y}e^{-yz}\mathrm{d}y.$

当$z\leqslant 0$时， $f_Z(z)=0$；

当$z>0$时，

$$f_Z(z)=\int_0^{+\infty}2y\mathrm{e}^{-2y}\mathrm{e}^{-yz}\mathrm{d}y=2\int_0^{+\infty}y\mathrm{e}^{-(2+z)y}\mathrm{d}y=\frac{2}{(2+z)^2}.$$

所以
$$f_Z(z)=\begin{cases}2/(2+z)^2, & z>0,\\ 0, & z\leqslant 0.\end{cases}$$

例 9 设某种型号的电子管的寿命(单位:h)近似地服从 $N(160,20^2)$分布,随机地选取 4 只,求其中没有一只的寿命小于 180 h 的概率.

解 以 $X_i(i=1,2,3,4)$记第 i 只电子管的寿命,$X_i\sim N(160,20^2)$,设 $A=\min\{X_1,X_2,X_3,X_4\}$,则

$$P\{A\geqslant 180\}=\prod_{i=1}^{4}P\{X_i\geqslant 180\}=\left[1-\Phi\left(\frac{180-160}{20}\right)\right]^4$$
$$=[1-\Phi(1)]^4=0.000634.$$

例 10 设某炮群向同一目标发射 n 发炮弹,炮弹的发射是独立的,每发炮弹射程的分布函数均为 $F(x)$. 求最大射程 U、最小射程 V 及(U,V)的联合分布函数.

解 以 X_i 记第 i 发炮弹的射程,显然 $X_1,X_2,\cdots,X_n$ 相互独立且同分布.

$$U=\max\{X_1,X_2,\cdots,X_n\},\quad V=\min\{X_1,X_2,\cdots,X_n\}.$$

$$\begin{aligned}F_U(u)&=P\{\max\{X_1,X_2,\cdots,X_n\}\leqslant u\}\\&=P\{X_1\leqslant u,X_2\leqslant u,\cdots,X_n\leqslant u\}\\&=P\{X_1\leqslant u\}P\{X_2\leqslant u\}\cdots P\{X_n\leqslant u\}=[F(u)]^n,\end{aligned}$$

$$\begin{aligned}F_V(v)&=P\{\min\{X_1,X_2,\cdots,X_n\}\leqslant v\}\\&=1-P\{\min\{X_1,X_2,\cdots,X_n\}>v\}\\&=1-P\{X_1>v\}P\{X_2>v\}\cdots P\{X_n>v\}=1-[F(v)]^n,\end{aligned}$$

$$\begin{aligned}F_{UV}(u,v)&=P\{\max\{X_1,X_2,\cdots,X_n\}\leqslant u,\min\{X_1,X_2,\cdots,X_n\}\leqslant v\}\\&=P\{\max\{X_1,X_2,\cdots,X_n\}\leqslant u\}\\&\quad-P\{\max\{X_1,X_2,\cdots,X_n\}\leqslant u,\min\{X_1,X_2,\cdots,X_n\}>v\}.\end{aligned}$$

因为 $P\{\max\{X_1,X_2,\cdots,X_n\}\leqslant u,\min\{X_1,X_2,\cdots,X_n\}>v\}$

$$=\begin{cases} P\{\varnothing\}=0, & u\leqslant v, \\ P\{v\leqslant X_1\leqslant u,\cdots,v<X_n\leqslant u\}, & u>v, \end{cases}$$

$$=\begin{cases} 0, & u\leqslant v, \\ [F(u)-F(v)]^n, & u>v, \end{cases}$$

所以 $F_{UV}(u,v)=\begin{cases} [F(u)]^n, & u\leqslant v, \\ [F(u)]^n-[F(u)-F(v)]^n, & u>v. \end{cases}$

例 11 设随机变量(X,Y)的分布律为

Y \ X	0	1	2	3	4	5
0	0	0.01	0.03	0.05	0.07	0.09
1	0.01	0.02	0.04	0.05	0.06	0.08
2	0.01	0.03	0.05	0.05	0.05	0.06
3	0.01	0.02	0.04	0.06	0.06	0.05

求:(1) $P\{X=2|Y=2\}$,$P\{Y=3|X=0\}$;

(2) $V=\max(X,Y)$的分布律;

(3) $U=\min(X,Y)$的分布律;

(4) $W=X+Y$ 的分布律.

解 先分别求出X 和Y 的边缘分布律,即

X	0	1	2	3	4	5
p_k	0.03	0.08	0.16	0.21	0.24	0.28

Y	0	1	2	3
p_k	0.25	0.26	0.25	0.24

(1) $P\{X=2|Y=2\}=P\{X=2,Y=2\}/P\{Y=2\}=2/5$,

$P\{Y=3|X=0\}=P\{X=0,Y=3\}/P\{X=0\}=1/3$.

(2) $P\{V=k\}=P\{X=k,Y=k\}+P\{X=k,Y<k\}$

$+P\{X<k,Y=k\}$,

故

V	0	1	2	3	4	5
p_k	0	0.04	0.16	0.28	0.24	0.28

(3) $P\{U=k\}=P\{X=k,Y=k\}+P\{X=k,Y>k\}$
$$+P\{X>k,Y=k\},$$

故

U	0	1	2	3
p_k	0.28	0.30	0.25	0.17

(4) 由
$$P\{W=X+Y=k\}$$
$$=\sum_{i=0}^{k}\{X=i,Y=k-i\}$$
$$=\sum_{i=0}^{k}P\{X=i\}P\{Y=k-i\},\ k=0,1,\cdots,8,$$

W 的分布律为

W	0	1	2	3	4	5	6	7	8
p_k	0	0.02	0.06	0.13	0.19	0.24	0.19	0.12	0.05

下面介绍怎样用引入随机变量函数组的方法求两个随机变量的函数的分布.

一般地,以所讨论的两个随机变量的函数作为函数组中的一个函数,以所给两个随机变量中的一个作为函数组中的另一个函数.实际上与微积分中二元函数的坐标变换是一致的,然后按前面介绍的步骤去做,即可得到我们需要求的结果.

例 12 设随机变量(X,Y)的概率密度为
$$f(x,y)=\begin{cases}3x, & 0<x<1,0<y<x,\\ 0, & \text{其它},\end{cases}$$
求随机变量 $Z=X-Y$ 的概率密度.

解一 用分布函数法.
$$F_Z(z)=\iint\limits_{x-y\leqslant z}f(x,y)\mathrm{d}x\mathrm{d}y=\int_0^1\mathrm{d}x\int_{x-z}^{+\infty}f(x,y)\mathrm{d}y.$$
由参变量积分的求导公式知,当$z<0$ 时,$y=x-u>x$,则
$$f_Z(z)=[F_Z(z)]'=\int_0^1\left[\int_{x-z}^{+\infty}f(x,y)\mathrm{d}y\right]_z'\mathrm{d}x=\int_0^1 f(x,z-x)\mathrm{d}x,$$

其中 $f(\alpha, x-z)=0$，所以 $f_Z(z)=0$.

当 $0\leqslant z<1$ 时，若 $z<x<1$，则 $f(x,z-x)=3x$；若 $0<x\leqslant z<1$，则 $f(x,x-u)=0$. 所以

$$f_Z(z)=\int_0^z 0\cdot \mathrm{d}x+\int_z^1 3x\mathrm{d}x=3(1-z^2)/2;$$

当 $z\geqslant 1$ 时，$f(x,x-z)=0$，　$f_Z(z)=0$.

所以
$$f_Z(z)=\begin{cases}3(1-z)^2/2, & 0\leqslant z<x<1,\\ 0, & \text{其它}.\end{cases}$$

解二　引入随机变量函数组

$$\begin{cases}U=X-Y,\\ V=Y,\end{cases}\Longrightarrow\begin{cases}X=U+V,\\ Y=V,\end{cases}$$

其雅可比行列式 $J=\dfrac{\partial(x,y)}{\partial(u,v)}=\begin{vmatrix}1 & 1\\ 0 & 1\end{vmatrix}=1$，所以 (U,V) 的联合密度为

$$f_{UV}(u,v)=f[x(u,v),y(u,v)]\cdot|J|=f[x(u,v),y(u,v)].$$

当 $0<x<1, 0<y<x$ 时，$u+v<1, 0<v<u+v$，所以 $f_{UV}(u,v)$ 的非零区域为 $u>0, v>0, u+v<1$. 于是

$$f_{UV}(u,v)=\begin{cases}3(u+v), & u>0,v>0,u+v<1,\\ 0, & \text{其它},\end{cases}$$

得
$$f_V(u)=f_{X+Y}(x+y)=\int_{-\infty}^{+\infty}f_{UV}(u,v)\mathrm{d}v$$
$$=\int_0^{1-u}3(u+v)\mathrm{d}v=3(1-u^2)/2,\quad 0<u<1.$$

与解一结果相同.

例 13　设电流强度 I 和电阻 R 为相互独立的随时间服从下列密度的随机变量：

$$f_I(i)=\begin{cases}6i(1-i), & 0<i<1,\\ 0, & \text{其它},\end{cases}\qquad f_R(r)=\begin{cases}2r, & 0<r<1,\\ 0, & \text{其它},\end{cases}$$

求功率 $W=IR^2$ 的概率密度.

解一　用分布函数法.

$$F_W(w)=P\{I^2R\leqslant w\}=\iint\limits_{i^2r\leqslant w} f_I(i)f_R(r)\mathrm{d}i\mathrm{d}r,$$

当$w\leqslant 0$时，$F_W(w)=0$，故 $f_W(w)=0$；

当$w\geqslant 1$时，$F_W(w)=1$，故 $f_W(w)=0$；

当$0<w<1$时，

$$F_W(w)=\int_0^1\mathrm{d}r\int_0^{\min(1,\sqrt{w/r})}2r\times 6i(1-i)\mathrm{d}i$$

$$=\int_0^w\mathrm{d}r\int_0^1 12ir(1-i)\mathrm{d}i+\int_w^1\mathrm{d}r\int_0^{\sqrt{w/r}}12ir(1-i)\mathrm{d}i.$$

由参变量积分的求导公式，得

$$f_W(w)=[F_W(w)]'$$

$$=12w\int_0^1 i(1-i)\mathrm{d}i-\int_0^{\sqrt{w/r}}12wi(1-i)\mathrm{d}i$$

$$+\int_w^1 12r\sqrt{\frac{w}{r}}\left(1-\sqrt{\frac{w}{r}}\right)\frac{1}{2}\times\frac{1}{\sqrt{rw}}\mathrm{d}r$$

$$=6[1-w-2\sqrt{w}(1-\sqrt{w})]$$

$$=6(1-\sqrt{w})^2.$$

解二 建立随机变量函数组

$$\begin{cases}W=I^2R,\\U=R,\end{cases}\Longrightarrow\begin{cases}R=U,\\I=\sqrt{W/U},\end{cases}$$

其雅可比行列式$J=\partial(r,i)/\partial(w,u)=1/(2\sqrt{wu})$.

随机变量(W,U)的概率密度为

$$f_{WU}(w,u)=f_I(\sqrt{w/u})f_R(u)\cdot|J|$$

$$=6\sqrt{w/u}(1-\sqrt{w/u})\cdot 2u/(2\sqrt{wu})$$

$$=6(1-\sqrt{w/u}),\ 0<w<u<1,$$

于是 $$f_W(w)=\int_{-\infty}^{+\infty}f_{WU}(w,u)\mathrm{d}u=\int_w^1 6(1-\sqrt{w/u})\mathrm{d}u$$

$$=6(1-\sqrt{w})^2.$$

例14 设随机变量X,Y相互独立，且都服从$N(0,1)$，令$Z=X$

$+Y,W=X-Y$，求(Z,W)的概率密度函数.

解 引入随机变量函数组

$$\begin{cases}Z=X+Y,\\ W=X-Y,\end{cases}\Longrightarrow\begin{cases}X=(Z+W)/2,\\ Y=(Z-W)/2.\end{cases}$$

其雅可比行列式$J=\partial(x,y)/\partial(z,w)=-1/2$，$|J|=1/2$，所以$(Z,W)$的概率密度为

$$f_{ZW}(z,w)=\frac{1}{2\pi}e^{-[(z+w)^2/4+(z-w)^2/4]}\times\frac{1}{2}=\frac{1}{4\pi}e^{-(z^2+w^2)/2}.$$

例 15 设随机变量(X,Y,Z)的概率密度为

$$f(x,y,z)=\begin{cases}e^{-(x+y+z)}, & x>0,y>0,z>0,\\ 0, & \text{其它},\end{cases}$$

求$U=(X+Y+Z)/3$的概率密度函数.

解 引入随机变量函数组

$$\begin{cases}U=(X+Y+Z)/3,\\ V=Y,\\ W=Z,\end{cases}\Longrightarrow\begin{cases}X=3U-V-W,\\ Y=V,\\ Z=W,\end{cases}$$

其雅可比行列式$J=\partial(x,y,z)/\partial(u,v,w)=3$. 所以，$(U,V,W)$的概率密度为

$$f(u,v,w)=\begin{cases}3e^{-3u}, & 3u-v-w>0,v>0,w>0,\\ 0, & \text{其它}.\end{cases}$$

又由$w>0,3u-v-w>0$得，$0<w<3u-v$，故

$$f_{UV}(u,v)=\begin{cases}\int_0^{3u-v}3e^{-3u}dw, & 3u-v>0,v>0,\\ 0, & \text{其它},\end{cases}$$

$$f_U(u)=\begin{cases}\int_0^{3u}dv\int_0^{3u-v}3e^{-3u}dw=\frac{27}{2}u^2e^{-3u}, & u>0,\\ 0, & \text{其它}.\end{cases}$$

例 16 设随机变量X和Y相互独立且同分布，密度函数

$$f_T(t)=\begin{cases}e^{-t}, & t>0,\\ 0, & t\leqslant 0,\end{cases}$$

证明:随机变量$U=X+Y$与随机变量$V=X/Y$相互独立.

解 引入随机变量函数组

$$\begin{cases}U=X+Y,\\V=X/Y,\end{cases}\Longrightarrow\begin{cases}X=UV/(V+1),\\Y=U/(V+1),\end{cases}\quad u>0,v>0,$$

其雅可比行列式$J=\partial(x,y)/\partial(u,v)=-u/(v+1)^2$. 所以,$(U,V)$的联合概率密度为

$$f_{UV}(u,v)=\mathrm{e}^{-(uv+u)/(v+1)}\cdot u/(v+1)^2,\ u>0,v>0.$$

即
$$f_{UV}(u,v)=\begin{cases}\dfrac{u\mathrm{e}^{-u}}{(v+1)^2}, & u>0,v>0,\\ 0, & \text{其它},\end{cases}$$

$$f_U(u)=\begin{cases}\displaystyle\int_0^{+\infty}\frac{u\mathrm{e}^{-u}}{(v+1)^2}\mathrm{d}v=u\mathrm{e}^{-u}, & u>0,\\ 0, & \text{其它},\end{cases}$$

$$f_V(v)=\begin{cases}\displaystyle\int_0^{+\infty}\frac{u\mathrm{e}^{-u}}{(v+1)^2}\mathrm{d}u=\frac{1}{(v+1)^2}, & v>0,\\ 0, & \text{其它}.\end{cases}$$

显然,$f_{UV}(u,v)=f_U(u)f_V(v)$,所以$X+Y$与X/Y相互独立.

硕士研究生入学试题分析

一、本章考试要求

1. 理解二维随机变量的概念;理解二维随机变量的联合分布的概念、性质及两种基本形式:(1)离散型联合概率分布、边缘分布和条件分布,(2)连续型联合概率密度、边缘密度和条件密度;会利用二维概率分布求有关事件的概率.

2. 理解随机变量的独立性及不相关的概念,掌握离散型和连续型随机变量独立的条件.

3. 掌握二维均匀分布,了解二维正态分布的概率密度,理解

其中参数的概率意义.

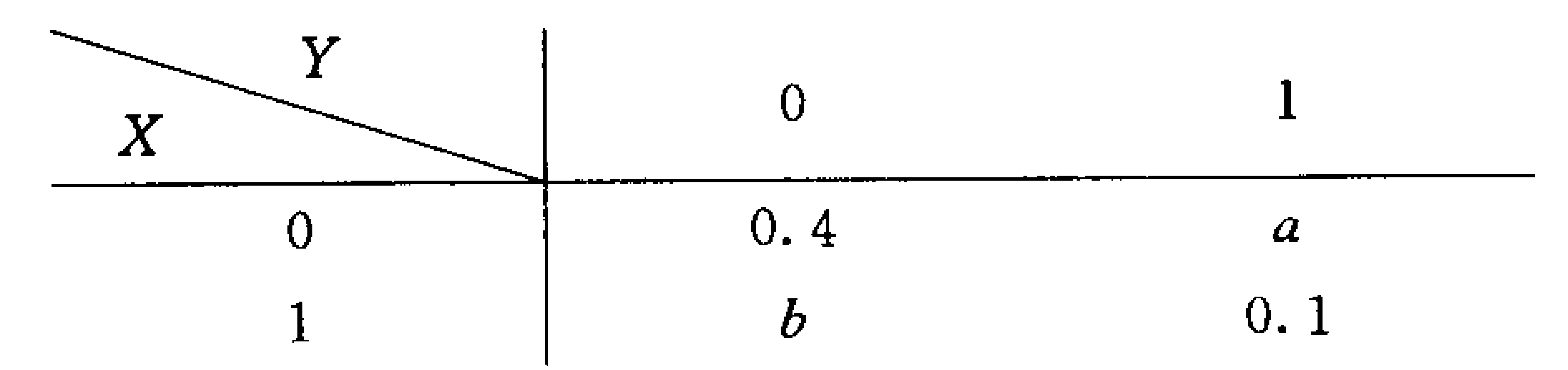
4. 会求两个独立随机变量的简单函数的分布.

二、本章重点内容

二维随机变量及其概率分布(包括联合分布、边缘分布和条件分布),相关事件的概率计算,两个独立随机变量的函数的分布及其概率的计算,综合题.

(一) 二维随机变量及其分布

1. 设二维随机变量(X,Y)的概率分布为



X \ Y	0	1
0	0.4	a
1	b	0.1

若随机事件$\{X=0\}$与$\{X+Y=1\}$相互独立,则(　　).

(A) $a=0.2,b=0.3$;　　(B) $a=0.1,b=0.4$;

(C) $a=0.3,b=0.2$;　　(D) $a=0.4,b=0.1$.

(2005 年一、三、四)

解 选(D). 设X与Y相互独立,则

$$P\{X=0\}P\{Y=1\}=a \Longrightarrow a=0.4 \text{ 或 } 0.1,$$

$$P\{X=1\}P\{Y=0\}=b \Longrightarrow b=0.1 \text{ 或 } 0.4,$$

当$a=0.4,b=0.1$时,

$$P\{X+Y=1\}P\{X=0\}=0.5\times 0.8=P\{X=0,X+Y=1\},$$

故$\{X=0\}$与$\{X+Y=1\}$相互独立;当$a=0.1,b=0.4$时,

$$P\{X+Y=1\}P\{X=0\}=0.5\times 0.5\neq P\{X=0,X+Y=1\},$$

即$\{X=0\}$与$\{X+Y=1\}$不相互独立.

2. 设二维随机变量(X,Y)的概率密度为

$$f(x,y)=\begin{cases}6x, & 0\leqslant x\leqslant y\leqslant 1,\\ 0, & \text{其它},\end{cases}$$

则$P\{X+Y\leqslant 1\}=$________.

(2003 年一)

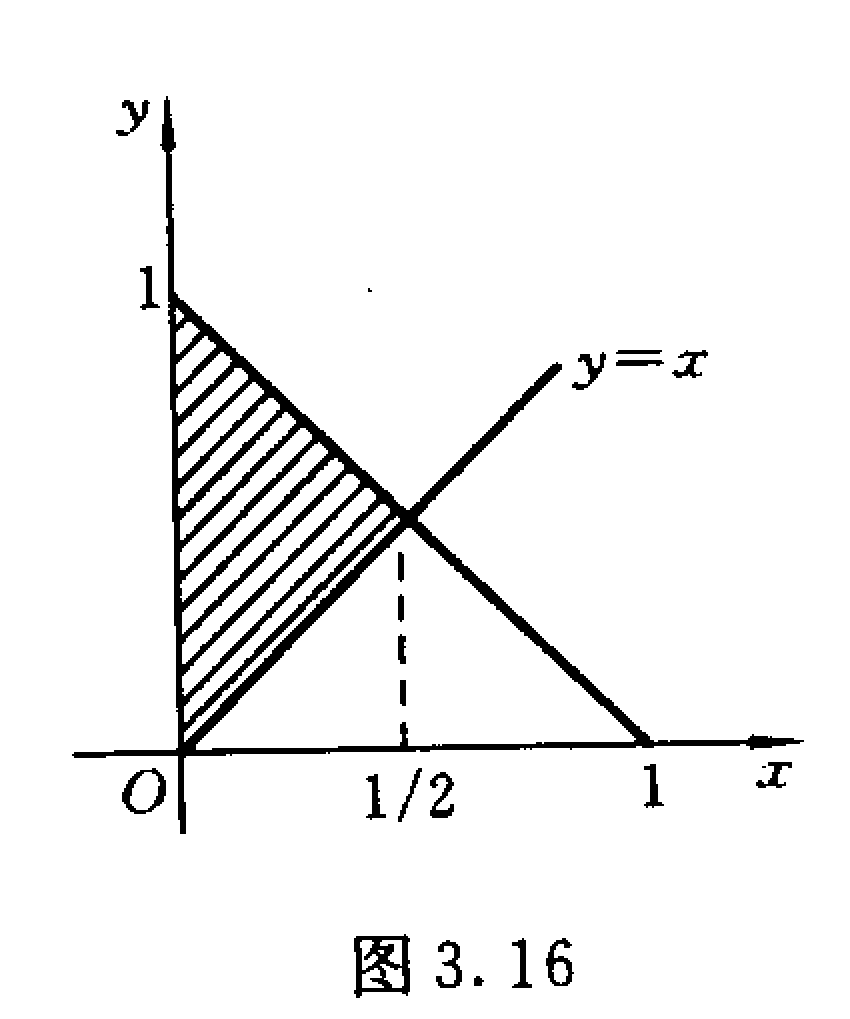


图 3.16

解 由 $x+y\leqslant 1$ 与 $x\leqslant y$ 知 D 如图 3.16所示，故

$$P\{X+Y\leqslant 1\}=\iint_D 6x\mathrm{d}x\mathrm{d}y$$
$$=\int_0^{1/2}6x\mathrm{d}x\int_x^{1-x}\mathrm{d}y$$
$$=\int_0^{1/2}6x(1-2x)\mathrm{d}x$$
$$=(3x^2-4x^3)\Big|_0^{1/2}=\frac{1}{4}.$$

3. 设随机变量

$$X_i\sim\begin{pmatrix}-1 & 0 & 1\\ 1/4 & 1/2 & 1/4\end{pmatrix},\quad i=1,2,$$

且满足 $P\{X_1X_2=0\}=1$，则 $P\{X_1=X_2\}=$(　　).

(A) 0;　(B) 1/4;　(C) 1/2;　(D) 1.　　　(1999 年三)

解 选(A). 由 $P\{X_1X_2=0\}=1$ 知

$$P\{-1,-1\}=P\{-1,1\}=P\{1,-1\}=P\{1,1\}=0,$$

从而知

$$P\{0,-1\}=P\{0,1\}=P\{-1,0\}=P\{1,0\}=1/4.$$

于是
$$P\{0,0\}=P\{X=0\}-P\{0,-1\}-P\{0,1\}$$
$$=1/2-1/4-1/4=0.$$

故 $P\{X_1=X_2\}=0$，X_1，X_2 的分布律为

X_2 \ X_1	-1	0	1	$p_{\cdot j}$
-1	0	1/4	0	1/4
0	1/4	0	1/4	1/2
1	0	1/4	0	1/4
$p_{i\cdot}$	1/4	1/2	1/4	1

4. 设两个随机变量 X 与 Y 相互独立且同分布，

$$P\{X=-1\}=P\{Y=-1\}=1/2,\quad P\{X=1\}=P\{Y=1\}=1/2,$$

则下列各式中成立的是(　　).

(A) $P\{X=Y\}=1/2$；　　(B) $P\{X=Y\}=1$；

(C) $P\{X+Y=0\}=1/4$；　　(D) $P\{XY=1\}=1/4$.

（1997 年三）

解　选(A). 因 X,Y 相互独立，由边缘分布可求得联合分布律为

$Y \backslash X$	-1	1	$p_{\cdot j}$
-1	1/4	1/4	1/2
1	1/4	1/4	1/2
$p_{i\cdot}$	1/2	1/2	1

故

$$P\{X=Y\}=1/2.$$

5. 设平面区域 D 由曲线 $y=1/x$ 及直线 $y=0, x=1, x=\mathrm{e}^2$ 所围成. 二维随机变量 (X,Y) 在 D 上服从均匀分布，则 (X,Y) 关于 X 的边缘概率密度在 $x=2$ 处的值为____.　　（1998 年一）

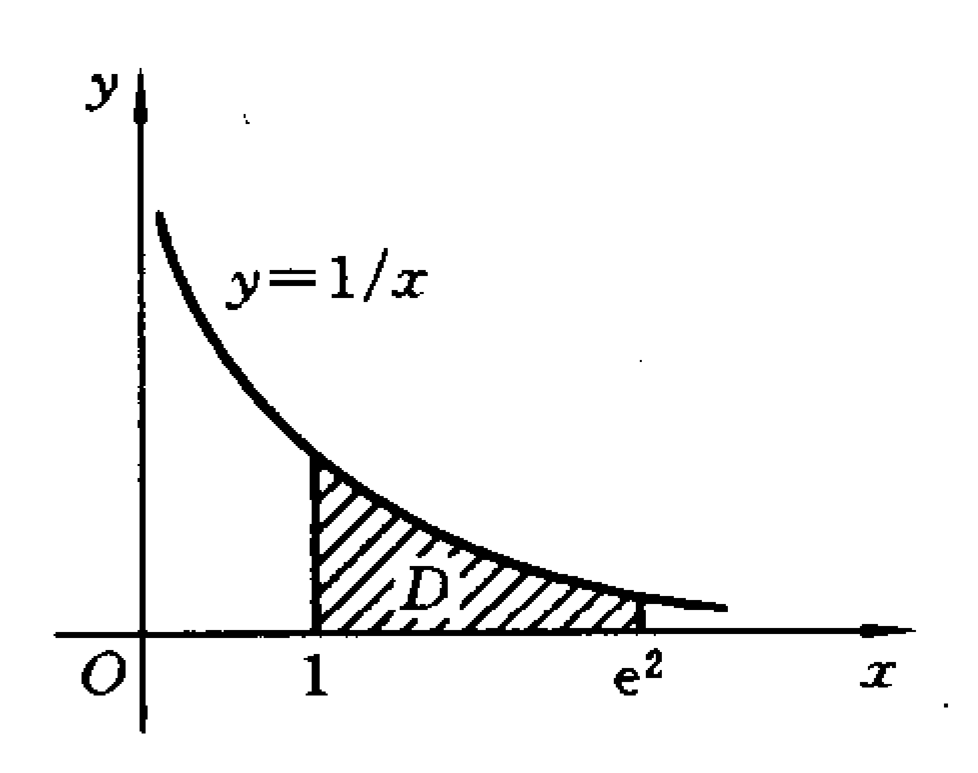
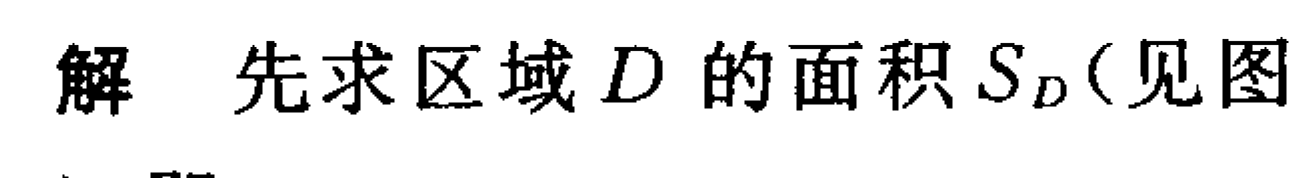


图 3.17



解　先求区域 D 的面积 S_D（见图 3.17），即

$$S_D=\int_1^{\mathrm{e}^2}\mathrm{d}x\int_0^{1/x}\mathrm{d}y=\int_1^{\mathrm{e}^2}\frac{1}{x}\mathrm{d}x=\ln x\Big|_1^{\mathrm{e}^2}=2,$$

故

$$f(x,y)=\begin{cases}1/2, & (x,y)\in D,\\ 0, & (x,y)\notin D.\end{cases}$$

$$f_X(x)=\int_0^{1/x}\frac{1}{2}\mathrm{d}x=\frac{1}{2}x,\ 1\leqslant x\leqslant \mathrm{e}^2.$$

于是

$$f_X(2)=1/4.$$

6. 设随机变量 X 和 Y 相互独立，二维随机变量 (X,Y) 的联合分布律及关于 X,Y 的边缘分布律的部分数值如下：

$X \backslash Y$	y_1	y_2	y_3	$p_{i\cdot}$
x_1		1/8		
x_2	1/8			
$p_{\cdot j}$	1/6			1

试将其余数值填入表中空白处.（1999 年一）

解 由 $$p_{\cdot 1}=1/6, p_{21}=1/8$$

知 $$p_{11}=1/24.$$

又由 $$p_{11}=p_{1\cdot}p_{\cdot 1}=p_{1\cdot}\times 1/6=1/24$$

知 $$p_{1\cdot}=1/4,\quad p_{2\cdot}=1-1/4=3/4,$$

$$p_{13}=1/4-1/24-1/8=1/12.$$

又由 $$p_{12}=p_{1\cdot}p_{\cdot 2}=1/4\times p_{\cdot 2}=1/8$$

知 $$p_{\cdot 2}=1/2,$$

故 $$p_{22}=1/2-1/8=3/8,\quad p_{\cdot 3}=1-1/6-1/2=1/3,$$

$$p_{23}=3/4\times 1/3=1/4.$$

因此

$X \backslash Y$	y_1	y_2	y_3	$p_{i\cdot}$
x_1	1/24		1/12	1/4
x_2		3/8	1/4	3/4
$p_{\cdot j}$		1/2	1/3	

7. 已知随机变量 X_1 和 X_2 的概率分布

$$X_1\sim\begin{pmatrix}-1 & 0 & 1\\ 1/4 & 1/2 & 1/4\end{pmatrix},\quad X_2\sim\begin{pmatrix}0 & 1\\ 1/2 & 1/2\end{pmatrix},$$

而且 $P\{X_1,X_2=0\}=1$.

（1）求 X_1 和 X_2 的联合分布；

（2）问：X_1 和 X_2 是否相互独立？为什么？（1999 年四）

解 （1）由 $P\{X_1,X_2=0\}=1$，可知

$$P\{X_1=-1,X_2=1\}=P\{X_1=1,X_2=1\}=0,$$

从而 $$P\{-1,0\}=P\{X_1=-1\}=1/4,$$

$$P\{0,1\}=P\{X_2=1\}=1/2,$$

$$P\{1,0\}=P\{X_1=1\}=1/4,$$

$$P\{0,0\}=1-1/4-1/2-1/4=0.$$

于是 X_1 和 X_2 的联合分布律为

X_2 \ X_1	-1	0	1	$p_{\cdot j}$
0	1/4	0	1/4	1/2
1	0	1/2	0	1/2
$p_{i\cdot}$	1/4	1/2	1/4	1

(2) 根据联合分布和边缘分布验证 $p_{ij}=p_{i\cdot}p_{\cdot j}$. 取 $p_{00}=0$，但 $p_{0\cdot}p_{\cdot 0}=1/2\times 1/2\neq 0$，故 X_1 与 X_2 不相互独立.

8. 已知随机变量 X 和 Y 的联合概率密度为

$$\varphi(x,y)=\begin{cases}4xy, & 0\leqslant x\leqslant 1,0\leqslant y\leqslant 1,\\ 0, & \text{其它},\end{cases}$$

求 X 和 Y 的联合分布函数 $F(x,y)$. (1995 年四)

解 (1) 对于 $x<0$ 或 $y<0$，有

$$F(x,y)=P\{X\leqslant x,Y\leqslant y\}=0.$$

(2) 对于 $0\leqslant x\leqslant 1,0\leqslant y\leqslant 1$，有

$$F(x,y)=4\int_0^x\int_0^y uv\,\mathrm{d}u\mathrm{d}v=x^2y^2.$$

(3) 对于 $x>1,y>1$，有 $F(x,y)=1$.

(4) 对于 $x>1,0\leqslant y\leqslant 1$，有

$$F(x,y)=P\{X\leqslant 1,Y\leqslant y\}=y^2.$$

(5) 对于 $y>1,0\leqslant x\leqslant 1$，有

$$F(x,y)=P\{X\leqslant x,Y\leqslant 1\}=x^2.$$

故 X 和 Y 的联合分布函数

$$F(x,y)=\begin{cases}0, & x<0\text{ 或 }y<0,\\ x^2y^2, & 0\leqslant x\leqslant 1,0\leqslant y\leqslant 1,\\ x^2, & 0\leqslant x\leqslant 1,1<y,\\ y^2, & 1<x,0\leqslant y\leqslant 1,\\ 1, & 1<x,1<y.\end{cases}$$

9. 已知随机变量 X 和 Y 的联合概率密度为

$$f(x,y)=\begin{cases}\mathrm{e}^{-(x+y)}, & 0<x<+\infty,0<y<+\infty,\\ 0, & \text{其它},\end{cases}$$

试求 $P\{X<Y\}$.（1989 年四）

解 因为 $x>0, y>0$，所以

$$P\{X<Y\}=\iint_{x<y} f(x,y)\mathrm{d}x\mathrm{d}y=\int_0^{+\infty}\left[\int_0^y \mathrm{e}^{-(x+y)}\mathrm{d}x\right]\mathrm{d}y$$

$$=\int_0^{+\infty}\mathrm{e}^{-y}\mathrm{d}y\int_0^y \mathrm{e}^{-x}\mathrm{d}x=\int_0^{+\infty}(\mathrm{e}^{-y}-\mathrm{e}^{-2y})\mathrm{d}y$$

$$=\left(-\mathrm{e}^{-y}+\frac{1}{2}\mathrm{e}^{-2y}\right)\Big|_0^{\infty}=\frac{1}{2}.$$

10. 甲、乙两人独立地进行两次射击，假设甲的命中率为 0.2，乙的命中率为 0.5，以 X 和 Y 分别表示甲和乙的命中次数，试求 X 和 Y 的联合分布律.（1990 年五）

解 以 $A_i(i=1,2)$ 记甲第 i 次射击命中，B_i 记乙第 i 次射击命中，射击相互独立，则 $P(A_i)=0.2, P(B_i)=0.5$. 设

$$X_i=\begin{cases}1, & A_i\text{ 发生},\\ 0, & A_i\text{ 不发生},\end{cases}\qquad Y_i=\begin{cases}1, & B_i\text{ 发生},\\ 0, & B_i\text{ 不发生},\end{cases}$$

于是

$$P\{X=0\}=P(\overline{A}_1,\overline{A}_2)=0.8^2=0.64,$$

$$P\{X=1\}=P(A_1,\overline{A}_2)+P(\overline{A}_1A_2)=2\times0.2\times0.8=0.32,$$

$$P\{X=2\}=P\{A_1A_2\}=0.2^2=0.04,$$

$$P\{Y=0\}=P(\overline{B}_1\overline{B}_2)=0.25,$$

$$P\{Y=1\}=2\times0.5^2=0.5,$$

$$P\{Y=2\}=P(B_1B_2)=0.25.$$

由 X,Y 的独立性可依 $p_{ij}=p_{i\cdot}\cdot p_{\cdot j}$ 写出 (X,Y) 的联合分布律如下：

Y \ X	0	1	2	$p_{i\cdot}$
0	0.16	0.08	0.01	0.25
1	0.32	0.16	0.02	0.5
2	0.16	0.08	0.01	0.25
$p_{\cdot j}$	0.64	0.32	0.04	1

11. 设某班车起点站上客人数 X 服从参数为 λ ($\lambda>0$) 的泊松分布，每位乘客在中途下车的概率为 p ($0<p<1$)，且中途下车与否相互独立，以 Y 表示在中途下车的人数，求：

(1) 在发车时有 n 个乘客的条件下，中途有 m 人下车的概率；

(2) 二维随机变量 (X,Y) 的概率分布. （2001 年一）

解 (1) $P\{Y=m|X=n\}=C_n^m p^m(1-p)^{n-m}$,

$$0\leqslant m\leqslant n,\quad n=0,1,2,\cdots,$$

(2) $P\{X=n,Y=m\}$

$$=P\{Y=m|X=n\}P\{X=n\}$$

$$=C_n^m p^m(1-p)^{n-m}\cdot\frac{e^{-\lambda}}{n!}\lambda^n,\quad 0\leqslant m\leqslant n,\quad n=0,1,2.$$

12. 设随机变量 X 在区间 $(0,1)$ 上服从均匀分布，在 $X=x$ $(0<x<1)$ 的条件下，随机变量 Y 在区间 $(0,x)$ 上服从均匀分布，求：

(1) 随机变量 X 和 Y 的联合概率密度；

(2) Y 的概率密度；

(3) 概率 $P\{X+Y>1\}$. （2004 年四）

解 (1) X 的概率密度为

$$f_X(x)=\begin{cases}1, & 0<x<1,\\ 0, & 其它,\end{cases}$$

在 $X=x$ $(0<x<1)$ 的条件下，Y 的条件概率密度为

$$f_{Y|X}(y|x)=\begin{cases}1/x, & 0<y<x,\\ 0, & 其它.\end{cases}$$

故当 $0<y<x<1$ 时，由 $f(x,y)=f_X(x)f_{Y|X}(y|x)$ 得出 X 与 Y 的联合概率密度如下：

$$f(x,y)=\begin{cases}1/x, & 0<y<x<1,\\ 0, & 其它.\end{cases}$$

(2) 由 $f_Y(y)=\int_{-\infty}^{+\infty}f(x,y)\mathrm{d}x$ 知，当 $0<y<1$ 时，

$$f_Y(y)=\int_y^1\frac{1}{x}\mathrm{d}x=-\ln y,$$

当 $y\leqslant 0$ 或 $y\geqslant 1$ 时，$f_Y(y)=0$. 即

$$f_Y(y)=\begin{cases}-\ln y, & 0<y<1,\\ 0, & 其它.\end{cases}$$

(3) $P\{X+Y>1\}=\iint\limits_{x+y>1} f(x,y)\mathrm{d}x\mathrm{d}y$

$$=\int_{1/2}^{1}\mathrm{d}x\int_{1-x}^{x}\frac{1}{x}\mathrm{d}y=\int_{1/2}^{1}\left(2-\frac{1}{x}\right)\mathrm{d}x$$

$$=1-\ln 2.$$

（二）两个随机变量的函数的分布

1. 设两个相互独立的随机变量 X 和 Y 分别服从正态分布 $N(0,1)$和$N(1,1)$，则(　　).

(A) $P\{X+Y\leqslant 0\}=1/2$；　(B) $P\{X+Y\leqslant 1\}=1/2$；

(C) $P\{X-Y\leqslant 0\}=1/2$；　(D) $P\{X-Y\leqslant 1\}=1/2$.

（1999 年一）

解 选(B). 因为$X\sim N(0,1)$，$Y\sim N(1,1)$，所以

$$X+Y\sim N(1,2),\quad (X+Y-1)/\sqrt{2}\sim N(0,1).$$

由正态分布的对称性知

$$P\{(X+Y-1)/\sqrt{2}\leqslant 0\}=1/2,$$

即$P\{X+Y-1\leqslant 0\}=1/2$，于是$P\{X+Y\leqslant 1\}=1/2$.

2. 设X 和Y 为两个随机变量，且

$$P\{X\geqslant 0,Y\geqslant 0\}=3/7,\quad P\{X\geqslant 0\}=P\{Y\geqslant 0\}=4/7,$$

则$P\{\max(X,Y)\geqslant 0\}=$________.　（1995 年一）

解
$$P\{\max(X,Y)\geqslant 0\}=P\{X\geqslant 0\}+P\{Y\geqslant 0\}-P\{X\geqslant 0,Y\geqslant 0\}$$
$$=4/7+4/7-3/7=5/7.$$

3. 设随机变量X 与Y 相互独立，X 的概率分布为

$$X\sim\begin{pmatrix}1 & 2\\ 0.3 & 0.7\end{pmatrix}$$

Y 的概率密度为$f(y)$，求随机变量$U=X+Y$ 的概率密度$g(u)$.

（2003 年三）

解 设Y 的分布函数为$F(y)$，则由全概率公式知，$U=X+Y$ 的分布函数为

$$
\begin{aligned}
G(u)&=P\{x+Y\leqslant u\}\\
&=0.3P\{X+Y\leqslant u\mid X=1\}+0.7\{X+Y\leqslant u\mid X=2\}\\
&=0.3P\{Y\leqslant u-1\mid X=1\}+0.7\{Y\leqslant u-2\mid X=2\}.
\end{aligned}
$$

因为X和Y相互独立,故

$$
\begin{aligned}
G(u)&=0.3P\{Y\leqslant u-1\}+0.7P\{Y\leqslant u-2\}\\
&=0.3F(u-1)+0.7F(u-2),
\end{aligned}
$$

从而,U的概率密度为

$$g(u)=G'(u)=0.3f(u-1)+0.7f(u-2).$$

4. 设二维随机变量(X,Y)的概率密度为

$$f(x,y)=\begin{cases}1, & 0<x<1,0<y<2x,\\ 0, & 其它,\end{cases}$$

求:(1) (X,Y)的边缘概率密度$f_X(x),f_Y(y)$;

(2) $Z=2X-Y$的概率密度$f_Z(z)$;

(3) $P\{Y<1/2\mid X\leqslant 1/2\}$.(数学一不考.)

(2005年一、三、四)

解 (1) 当$0<x<1$时,

$$f_X(x)=\int_{-\infty}^{+\infty}f(x,y)\mathrm{d}y=\int_0^{2x}\mathrm{d}y=2x;$$

当$x\leqslant 0$或$x\geqslant 1$时,$f_X(x)=0$.故有

$$f_X(x)=\begin{cases}2x, & 0<x<1,\\ 0, & 其它.\end{cases}$$

当$0<y<2$时,

$$f_Y(y)=\int_{-\infty}^{+\infty}f(x,y)\mathrm{d}x=\int_{y/2}^{1}\mathrm{d}x=1-y/2;$$

当$y\leqslant 0$或$y\geqslant 2$时,$f_Y(y)=0$.故有

$$f_Y(y)=\begin{cases}1-y/2, & 0<y<2,\\ 0, & 其它.\end{cases}$$

(2) 当$z\leqslant 0$时, $F_Z(z)=0$;

当$z\geqslant 2$时, $F_Z(z)=1$;

当$0<z<2$时，

$$F_Z(z)=P\{2x-Y\leqslant z\}=\iint\limits_{2x-y\leqslant z}f(x,y)\mathrm{d}x\mathrm{d}y=z-\frac{z^2}{4}.$$

故 $$f_Z(z)=F'_Z(z)=\begin{cases}1-z/2, & 0<z<2,\\ 0, & \text{其它}.\end{cases}$$

(3) $P\{Y\leqslant 1/2\mid X\leqslant 1/2\}=\dfrac{P\{X\leqslant 1/2,Y\leqslant 1/2\}}{P\{X\leqslant 1/2\}}=\dfrac{3/16}{1/4}=\dfrac{3}{4}.$

5. 设二维随机变量(X,Y)在矩形域$G=\{x,y\mid 0\leqslant x\leqslant 2,0\leqslant y\leqslant 1\}$上服从均匀分布，试求边长为$X$和$Y$的矩形面积$S$的概率密度$f(s)$. (1999年四)

解 G的面积为2，(X,Y)的概率密度为

$$f(x,y)=\begin{cases}1/2, & x,y\in G,\\ 0, & \text{其它}.\end{cases}$$

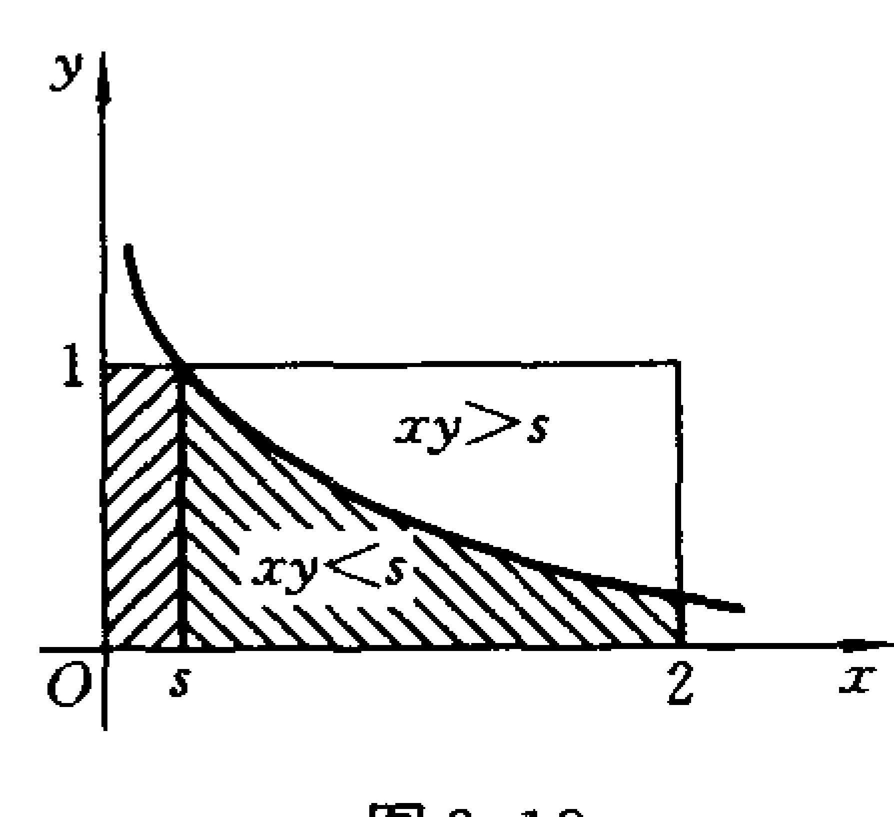


图3.18

以$F(s)=P\{S\leqslant s\}$表示S的分布函数，则当$s\leqslant 0$时，$F(s)=0$；当$s\geqslant 2$时，$F(s)=1$. 现考虑$0<s<2$的情形，由图3.18知

$$\begin{aligned}F(s)&=P\{S\leqslant s\}=P\{XY\leqslant s\}\\&=1-P\{XY>s\}=1-\iint\limits_{xy>s}\frac{1}{2}\mathrm{d}x\mathrm{d}y\\&=1-\frac{1}{2}\int_s^2\mathrm{d}x\int_{s/x}^1\mathrm{d}y\\&=s(1+\ln 2-\ln s)/2.\end{aligned}$$

于是 $$g(s)=F'(s)=\begin{cases}(\ln 2-\ln s)/2, & \text{当}0<s<2,\\ 0, & \text{其它}.\end{cases}$$

6. 假设随机变量X_1,X_2,X_3,X_4相互独立且同分布，$P\{X_i=0\}=0.6$，$P\{X_i=1\}=0.4$，$i=1,2,3,4$，求行列式

$$X=\begin{vmatrix}X_1 & X_2\\ X_3 & X_4\end{vmatrix}$$

的概率分布. (1994年四)

解 设$Y_1=X_1X_4$，$Y_2=X_2X_3$，则$X=Y_1-Y_2$，随机变量Y_1,Y_2相

互独立且同分布，有

$$P\{Y_1=1\}=P\{Y_2=1\}=P\{X_2=1,X_3=1\}=0.16,$$

$$P\{Y_1=0\}=P\{Y_2=0\}=1-0.16=0.84.$$

随机变量$X=Y_1-Y_2$有三个可能值$-1,0,1$，易见

$$P\{X=-1\}=P\{Y_1=0,Y_2=1\}=0.84\times 0.16=0.1344,$$

$$P\{X=1\}=P\{Y_1=1,Y_2=0\}=0.16\times 0.84=0.1344,$$

$$P\{X=0\}=1-P\{X=-1\}-P\{X=1\}=0.7312.$$

于是行列式的概率分布为

$$X=\begin{vmatrix} X_1 & X_2 \\ X_3 & X_4 \end{vmatrix}\sim\begin{pmatrix} -1 & 0 & 1 \\ 0.1344 & 0.7312 & 0.1344 \end{pmatrix}.$$

7. 设随机变量X与Y相互独立，X服从正态分布$N(\mu,\sigma^2)$，Y服从$[-\pi,\pi]$上的均匀分布，试求$Z=X+Y$的概率密度函数(计算结果用标准正态分布函数Φ来表示)，其中

$$\Phi(x)=\int_{-\infty}^{x}\frac{1}{\sqrt{2\pi}}e^{-t^2/2}dt. \qquad (1992\text{ 年一})$$

解 因为$f_X(x)=\dfrac{1}{\sqrt{2\pi}\sigma}e^{-(x-\mu)^2/(2\sigma^2)}$，$-\infty<x<+\infty$；

$$f_Y(y)=\begin{cases}\dfrac{1}{2\pi}, & -\pi<y<\pi, \\ 0, & \text{其它}, \end{cases}$$

X,Y相互独立，由卷积公式，有

$$f_Z(z)=\int_{-\infty}^{+\infty}f_X(z-y)f_Y(y)dy=\int_{-\pi}^{\pi}\frac{1}{\sqrt{2\pi}\sigma}e^{-(z-y-\mu)^2/(2\sigma^2)}\cdot\frac{1}{2\pi}dy$$

$$=\frac{1}{2\pi}\cdot\frac{1}{\sqrt{2\pi}\sigma}\int_{-\pi}^{\pi}e^{-(z-y-\mu)^2/(2\sigma^2)}dy \ (\text{令 } t=z-y-\mu)$$

$$=\frac{-1}{(2\pi)^{3/2}}\int_{(z+\pi-\mu)/\sigma}^{(z-\pi-\mu)/\sigma}e^{-t^2/2}dt$$

$$=\frac{1}{2\pi}\int_{-\infty}^{(z+\pi-\mu)/\sigma}\frac{1}{\sqrt{2\pi}}e^{-t^2/2}dt-\frac{1}{2\pi}\int_{-\infty}^{(z-\pi-\mu)\sigma}\frac{1}{\sqrt{2\pi}}e^{-t^2/2}dt$$

$$=\frac{1}{2\pi}\left[\Phi\left(\frac{z+\pi-\mu}{\sigma}\right)-\Phi\left(\frac{z-\pi-\mu}{\sigma}\right)\right].$$

8. 设二维随机变量(X,Y)的概率密度为

$$f(x,y)=\begin{cases}2e^{-(x+2y)}, & x>0,y>0,\\ 0, & 其它,\end{cases}$$

求随机变量$Z=X+2Y$的分布函数. (1991 年一)

解 $F_Z(z)=P\{Z\geqslant z\}=P\{X+2Y\leqslant z\}=\iint\limits_{x+2y\leqslant z} f(x,y)\mathrm{d}x\mathrm{d}y.$

当$z>0$时,

$$F_Z(z)=\int_0^z \mathrm{d}x\int_0^{(z-x)/2} 2e^{-(x+2y)}\mathrm{d}y=\int_0^z e^{-x}\mathrm{d}x\int_0^{(z-x)/2} 2e^{-2y}\mathrm{d}y$$

$$=\int_0^z (e^{-x}-e^{-z})\mathrm{d}x=1-e^{-z}-ze^{-z};$$

当$z\leqslant 0$时, $F_Z(z)=0.$

故
$$F_Z(z)=\begin{cases}1-e^{-z}-ze^{-z}, & z>0,\\ 0, & z\leqslant 0.\end{cases}$$

9. 一电子仪器由两部件构成,以X和Y分别表示部件的寿命(单位:kh),已知X和Y的联合分布函数为

$$F(x,y)=\begin{cases}1-e^{-0.5x}-e^{-0.5y}+e^{-0.5(x+y)}, & x\geqslant 0,y\geqslant 0,\\ 0, & 其它.\end{cases}$$

(1) 问X和Y是否相互独立;

(2) 求两个部件的寿命都超过100 h 的概率α. (1990 年四)

解 $F_X(x)=\lim\limits_{y\to+\infty} F(x,y)=\begin{cases}1-e^{-0.5x}, & x\geqslant 0,\\ 0, & x<0,\end{cases}$

$$F_Y(y)=\lim_{x\to+\infty} F(x,y)=\begin{cases}1-e^{-0.5y}, & y\geqslant 0,\\ 0, & y<0.\end{cases}$$

因为$F(x,y)=F_X(x)F_Y(y)$,所以X和Y相互独立.

$$\begin{aligned}\alpha&=P\{X>0.1\}P\{Y>0.1\}\\&=[1-P\{X<0.1\}][1-P\{Y<0.1\}]\\&=e^{-0.05}\cdot e^{-0.05}=e^{-0.1}.\end{aligned}$$

10. 设随机变量X,Y相互独立,其概率密度为

$$f_X(x)=\begin{cases}1, & 0\leqslant x\leqslant 1,\\ 0, & 其它,\end{cases}$$

$$f_Y(y)=\begin{cases}e^{-y}, & y>0,\\ 0, & y\geqslant 0,\end{cases}$$

求$Z=2X+Y$的概率密度函数.　(1987年一)

解　如图3.19所示，因为X,Y相互独立，所以有卷积公式

$$f_Z(z)=\int_{-\infty}^{+\infty}f_X(x)f_Y(z-2x)\mathrm{d}x,$$

图3.19

$$0\leqslant x\leqslant 1,z-2x>0.$$

当$z\leqslant 0$时，　$f_Z(z)=0$;

当$0<z\leqslant 2$时，

$$f_Z(z)=\int_0^{z/2}e^{-(z-2x)}\mathrm{d}x=e^{-z}\int_0^{z/2}e^{2x}\mathrm{d}x$$

$$=\frac{1}{2}e^{-z}(e^z-1)=\frac{1}{2}(1-e^{-z});$$

当$z>2$时，

$$f_Z(z)=\int_0^1 e^{-(z-2x)}\mathrm{d}x=e^{-z}\int_0^1 e^{2x}\mathrm{d}x=\frac{1}{2}e^{-z}(e^2-1).$$

故

$$f_Z(z)=\begin{cases}0, & z\leqslant 0,\\ (1-e^{-z})/2, & 0<z\leqslant 2,\\ e^{-z}(e^2-1)/2, & z>2.\end{cases}$$

11. 某仪器装有三只独立工作的同型号电子元件，其寿命(单位：h)都服从同一指数分布，分布密度为

$$f(x)=\begin{cases}\frac{1}{600}e^{-x/600}, & x>0,\\ 0, & x\leqslant 0.\end{cases}$$

求在仪器使用的最初200 h内至少有一只电子元件损坏的概率.

(1989年四、五)

解　设损坏只数是随机变量Y，$Y\sim B(3,p)$，A_i为最初200 h内第i只元件损坏事件，X_i为第i只元件的使用寿命，X_i相互独立且同

分布，故

$$P(A_i)=P\{X_i\leqslant 200\}=\int_1^{200}\frac{1}{600}\mathrm{e}^{-x/600}\mathrm{d}x$$

$$=-\mathrm{e}^{-x/600}\Big|_1^{200}=1-\mathrm{e}^{-1/3}\quad(i=1,2,3),$$

$$P\{Y\geqslant 1\}=1-(1-p)^3=1-(\mathrm{e}^{-1/3})^3=1-1/\mathrm{e}.$$

12. 设随机变量 X 与 Y 相互独立，且 X 服从均值为 1，标准差为 $\sqrt{2}$ 的正态分布，而 Y 服从标准正态分布，试求随机变量 $Z=2X-Y+3$ 的概率密度函数. (1989 年一)

解 $X\sim N(1,2)$，$Y\sim N(0,1)$，则由正态分布的参数可加性 $2X+Y\sim N(2,9)$，故 $Z\sim N(5,9)$，于是

$$f(z)=\frac{1}{3\sqrt{2\pi}}\mathrm{e}^{-(x-5)^2/18},\ -\infty<z<+\infty.$$

13. 设相互独立的两个随机变量 X,Y 具有同一分布律，且 X 的分布律为

X	0	1
p_k	1/2	1/2

则随机变量 $Z=\max\{X,Y\}$ 的分布律为________. (1994 年一)

解 因为 Z 只有两个值 0，1，所以

$$P\{Z=0\}=P\{X=0,Y=0\}=1/2\times 1/2=1/4,$$

$$P\{Z=1\}=P\{X=1,Y=0\}+P\{X=1,Y=1\}$$

$$=P\{X=0,Y=1\}$$

$$=1/2\times 1/2+1/2\times 1/2+1/2\times 1/2=3/4.$$

所以，随机变量 Z 的分布律为

Z	0	1
p_k	1/4	3/4

14. 设随机变量 X 的概率密度为 $f(x)=\mathrm{e}^{-x/2}$，$-\infty<x<+\infty$，问：随机变量 X 与 $|X|$ 是否相互独立？为什么？ (1993 年一)

解 $|X|$与X不相互独立.

因为事件$\{X<a\}\supset$事件$\{|X|<a\}$,所以

$$P\{X<a,|X|<a\}=P\{\{X<a\}\cap\{|X|<a\}\}$$
$$=P\{|X|<a\},$$

又 $\qquad P\{X<a\}<1,\quad P\{|X|<a\}>0,$

因而 $\qquad P\{X<a\}P\{|X|<a\}<P\{|X|<a\},$

即 $\qquad P\{X<a,|X|<a\}\neq P\{X<a\}P\{|X|<a\}.$

所以,$|X|$与X不相互独立.

15. 设随机变量X和Y的联合分布是正方形$G=\{(x,y):1\leqslant x\leqslant 3,1\leqslant y\leqslant 3\}$上的均匀分布,试求随机变量$U=|X-Y|$的概率密度$p(u)$. (2001 年三)

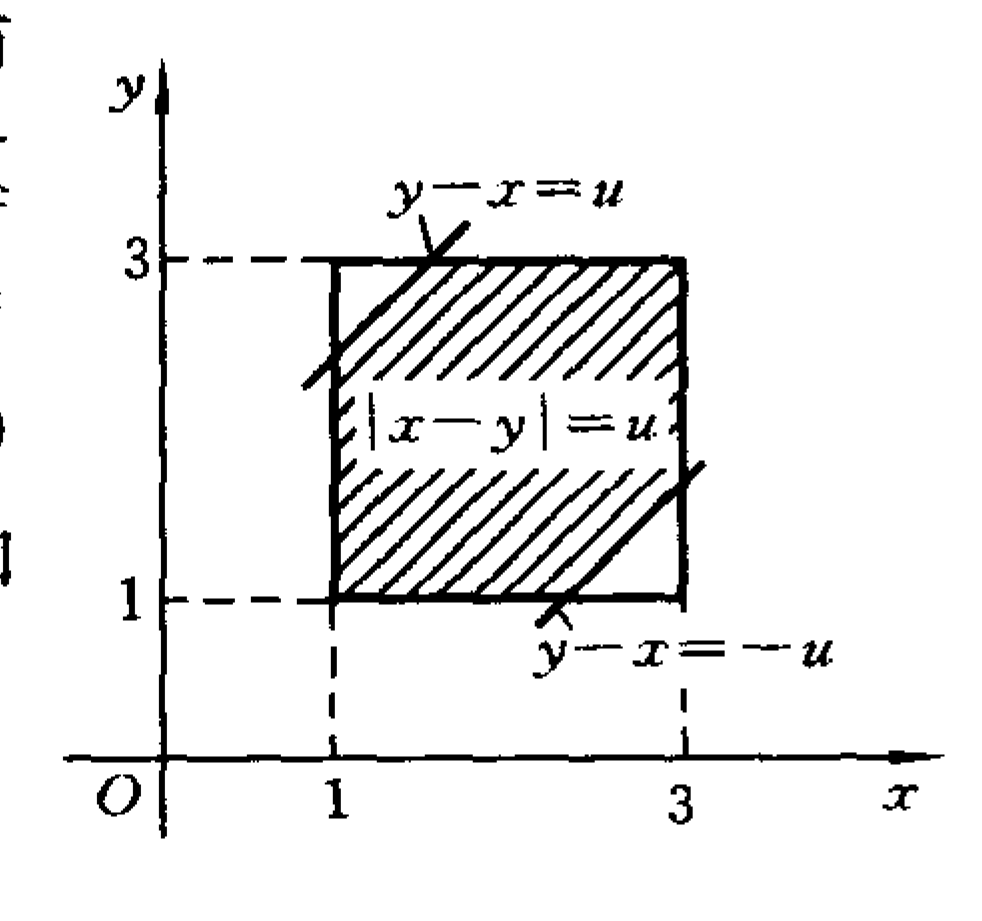


图3.20

解 如图3.20所示,由条件知X和Y的联合密度为

$$f(x,y)=\begin{cases}1/4, & 1\leqslant x\leqslant 3,1\leqslant y\leqslant 3,\\ 0, & \text{其它}.\end{cases}$$

以$F(u)=P\{U\leqslant u\}\ (-\infty<u<+\infty)$表示随机变量$U$的分布函数,显然,当$u\leqslant 0$时,$F(u)=0$;当$u\geqslant 2$时,$F(u)=1$.

设$0<u<2$,则

$$F(u)=\iint\limits_{|x-y|\leqslant u}f(x,y)\mathrm{d}x\mathrm{d}y=\iint\limits_{|x-y|\leqslant u}\frac{1}{4}\mathrm{d}x\mathrm{d}y$$
$$=\frac{1}{4}\times[4-(2-u)^2]=1-\frac{(2-u)^2}{4},$$

于是,随机变量U的密度为

$$p(u)=\begin{cases}(2-u)/2, & 0<u<2,\\ 0, & \text{其它}.\end{cases}$$

16. 假设一台设备开机后无故障工作的时间X服从指数分布,平均无故障工作时间$E(X)$为5 h. 设备定时开机,出现故障时自动关机,

而在无故障情况下工作 2 h 便关机. 试求该设备每次开机无故障工作的时间 Y 的分布函数 $F(y)$. （2002 年三、四）

解 设 X 的分布参数为 λ. 由于 $E(X)=1/\lambda=5$，可知 $\lambda=1/5$. 显然 $Y=\min\{X,2\}$.

对于 $y<0$， $F(y)=0$；

对于 $y\geqslant 2$， $F(y)=1$；

对于 $0\leqslant y<2$，

$$\begin{aligned}F(y)&=P\{Y\leqslant y\}=P\{\min\{X,2\}\leqslant y\}\\&=P\{X\leqslant y\}=1-e^{-y/5}.\end{aligned}$$

于是，Y 的分布函数为

$$F(y)=\begin{cases}0, & y<0,\\ 1-e^{-y/5}, & 0\leqslant y<2,\\ 1, & y\geqslant 2.\end{cases}$$

第四章　随机变量的数字特征

第一节　随机变量的数学期望与方差

主 要 内 容

一、数学期望

1. 离散型随机变量的数学期望

设离散型随机变量 X 的分布律为 $P\{X=x_k\}=p_k\quad(k=0,1,2,\cdots)$，若级数 $\sum_{k=1}^{\infty}x_kp_k$ 绝对收敛，则称此级数的和为随机变量 X 的数学期望或均值，记为 $E(X)$，即 $E(X)=\sum_{k=1}^{\infty}x_kp_k$.

2. 连续型随机变量的数学期望

设连续型随机变量 X 的概率密度为 $f(x)$，若积分 $\int_{-\infty}^{+\infty}xf(x)\mathrm{d}x$ 绝对收敛，则称此积分的值为随机变量 X 的数学期望，记为 $E(X)=\int_{-\infty}^{+\infty}xf(x)\mathrm{d}x$.

3. 随机变量函数的数学期望

设 $Y=g(X)$ 是随机变量 X 的函数，有：

(1) 若 X 是离散型的随机变量，分布律为 $P\{X=x_k\}=p_k\quad(k=1,2,\cdots)$，若 $\sum_{k=1}^{\infty}g(x_k)p_k$ 绝对收敛，则

$$E(Y)=E[g(X)]=\sum_{k=1}^{\infty}g(x_k)p_k.$$

(2) 若 X 是连续型的随机变量,概率密度为 $f(x)$,若积分 $\int_{-\infty}^{+\infty} g(x)f(x)\mathrm{d}x$ 绝对收敛,则

$$E(Y)=E[g(X)]=\int_{-\infty}^{+\infty} g(x)f(x)\mathrm{d}x \quad (g(x)\text{连续}).$$

若 $Z=g(X,Y)$ 是随机变量 X 和 Y 的函数,则有以下(3)、(4).

(3) 若 (X,Y) 是离散型随机变量,其分布律为

$$P\{X=x_i,Y=y_j\}=p_{ij}\ (i,j=1,2,\cdots),$$

则

$$E(Z)=E[g(X,Y)]=\sum_{i=1}^{\infty}\sum_{j=1}^{\infty} g(x_i,y_j)p_{ij}.$$

(4) 若 (X,Y) 是连续型随机变量,其密度函数为 $f(x,y)$,则

$$\begin{aligned}E(Z)&=E[g(X,Y)]\\&=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} g(x,y)f(x,y)\mathrm{d}x\mathrm{d}y \quad (g(x,y)\text{连续}).\end{aligned}$$

(3)和(4)中的级数与积分必须绝对收敛.

4. 数学期望的性质

(1) $E(C)=C$ (C 为常数).

(2) $E(CX)=CE(X)$ (C 为常数).

(3) 对两个随机变量 X 和 Y,有

$$E(X+Y)=E(X)+E(Y);$$

推广到多个随机变量和的情形,有

$$E(X_1+X_2+\cdots+X_n)=E(X_1)+E(X_2)+\cdots+E(X_n).$$

(4) 对两个相互独立的随机变量,有

$$E(XY)=E(X)E(Y);$$

推广到有限个相互独立的随机变量情形,有

$$E(X_1X_2\cdots X_n)=E(X_1)E(X_2)\cdots E(X_n).$$

二、方差

1. 方差的定义

设 X 为随机变量,若 $E\{[X-E(X)]^2\}$ 存在,则称 $E\{[X-E(X)]^2\}$ 为 X 的方差,记为

$$D(X)=\mathrm{var}(X)=E\{[X-E(X)]^2\}.$$

将 $\sqrt{D(X)}=\sigma(X)$ 称为随机变量 X 的均方差(标准差).

2. 方差的计算

(1) 若 X 为离散型随机变量,则

$$D(X)=\sum_{k=1}^{\infty}[x_k-E(X)]^2p_k.$$

(2) 若 X 是连续型随机变量,则

$$D(X)=\int_{-\infty}^{+\infty}[x-E(X)]^2f(x)\mathrm{d}x.$$

(3) 方差 $D(X)$ 的简捷计算公式为 $D(X)=E(X^2)-[E(X)]^2$.

3. 方差的基本性质

(1) 若 C 为常数,则 $D(C)=0$;

(2) 若 C 为常数,则 $D(CX)=C^2D(X)$;

(3) 若 X 与 Y 相互独立,则 $D(X+Y)=D(X)+D(Y)$(同时,$D(X-Y)=D(X)-D(Y)$);

(4) $D(X)=0\Longleftrightarrow P\{X=C\}=1$.

三、一些常用分布的数学期望与方差

$X\sim B(n,p)$,则 $E(X)=np$,$D(X)=np(1-p)$.

$X\sim\pi(\lambda)(P(\lambda))$,则 $E(X)=\lambda$,$D(X)=\lambda$.

$X\sim U(a,b)$,则 $E(X)=(a+b)/2$,$D(X)=(b-a)^2/12$.

$X\sim \mathrm{e}(\lambda)$,则 $E(X)=1/\lambda$,$D(X)=1/\lambda^2$.

$X\sim N(\mu,\sigma^2)$,则 $E(X)=\mu$,$D(X)=\sigma^2$.

疑难解析

1. 随机变量的数字特征在概率论中有什么实际意义?

答 要全面掌握一个随机变量的统计规律性,必须了解这个随机变量的分布函数.但是,在实际问题中要求得一个随机变量的分布函数是困难的,也往往是不必要的,通常只需要了解随机变量的某几

个数量指标就可以了. 这些指标就是概率论中的随机变量的数字特征. 一来,它们比较简单易求(在实际问题中可以通过样本进行估计);二来,它们已经足够满足解决实际问题的需要,并且刻画了随机变量的某些特征. 随机变量的数字特征在概率论与数理统计中有着广泛的应用.

2. 在数学期望定义中为什么要求级数 $\sum_{k=1}^{\infty} x_k p_k$ 和积分 $\int_{-\infty}^{+\infty} f(x)\mathrm{d}x$ 绝对收敛?

答 首先,随机变量的取值是随机的,不一定是按次序的,因此,在求和的时候,可能要改变项的次序. 其次,由高等数学知识可知,当级数绝对收敛时,不因改变项的次序而改变级数的和,因而期望值唯一存在. 积分是一个积分和式(也可以视为一个级数)的极限,因而也要求绝对收敛.

3. 为什么不用 $E[X-E(X)]$ 或 $E[|X-E(X)|]$ 代替方差作为衡量 X 取值偏离程度的一个尺度?

答 $X-E(X)$ 虽然反映了随机变量 X 的取值与平均值 $E(X)$ 的偏差,但偏差可能是正的也可能是负的,因而在求和时容易正、负抵消,故 $E[X-E(X)]$ 的值不能正确反映 X 取值与平均值的偏离程度.

对 $E[|X-E(X)|]$ 来说,虽然不再存在正、负抵消的问题,但是要取绝对值,给计算制造了一定的麻烦,所以 $E[|X-E(X)|]$ 也不适合作为衡量随机变量 X 取值偏离程度的一个尺度.

4. 随机变量 X 的数学期望 $E(X)$ 与方差 $D(X)$ 之间有什么联系?

答 (1) 若 $E(X)$ 存在,$D(X)$ 不一定存在.

如对二维随机变量 (X,Y),若概率密度为

$$f(x,y)=1/[\pi(x^2+y^2+1)^2],\quad -\infty<x,y<+\infty,$$

有 $E(X)=0, E(Y)=0$,但 $D(X)=+\infty, D(Y)=+\infty$.

(2) 若 $E(X)$ 不存在,则 $D(X)$ 一定不存在.

因为 $D(X)=E(X^2)-[E(X)]^2$,所以 $E(X)$ 不存在,必有 $D(X)$

不存在（如 X 服从柯西分布，$f(x)=1/[\pi(1+x^2)]$，$-\infty<x<+\infty$，$E(X)$不存在）.

(3) 若$D(X)$存在，则对任意常数C，有$D(X)\leqslant E(X-C)^2$，当且仅当$C=E(X)$时等号成立.

方法、技巧与典型例题分析

在计算随机变量的数字特征时，必须先分析已知的条件，根据不同的条件寻找不同的计算方法. 当分布已知时，一般只需依公式计算随机变量的数字特征. 所要注意的是要验证级数或积分的绝对收敛性，并尽量利用级数求和的技巧和积分的性质. 当分布未知时，可以先求出分布，再计算随机变量的数字特征. 但这比较麻烦，一般利用关于数字特征的定理和数字特征之间的关系来计算.

一、分布已知时，求数学期望与方差

例 1 设随机变量 X 的分布律如下：

X	-2	0	2
p_k	0.4	0.3	0.3

求：$E(X)$，$E(X^2)$，$E(3X^2+5)$，$D(X)$.

解 $E(X)=-2\times0.4+0\times0.3+2\times0.3=-0.2$，

$$E(X^2)=(-2)^2\times0.4+0^2\times0.3+2^2\times0.3=2.8,$$

$$E(3X^2+5)=3E(X^2)+5=3\times2.8+5=13.4,$$

$$D(X)=E(X^2)-[E(X)]^2=2.8-(-0.2)^2=2.76.$$

例 2 设随机变量 X 的分布律为

$$P\{X=k\}=\alpha^k/(1+\alpha)^{k+1},\ \alpha>0,k=0,1,\cdots,$$

求：$E(X)$和$D(X)$.

解 $E(X)=\sum\limits_{k=0}^{\infty}\dfrac{k\alpha^k}{(1+\alpha)^{k+1}}$，利用幂级数求和技巧，有

$$\sum_{k=1}^{\infty}kx^{k-1}=\sum_{k=1}^{\infty}(x^k)'=\left(\sum_{k=1}^{\infty}x^k\right)'=\left(\frac{1}{1-x}-1\right)'=\frac{1}{(1-x)^2}.$$

$$E(X)=\frac{\alpha}{(1+\alpha)^2}\sum_{k=1}^{\infty}k\left(\frac{\alpha}{1+\alpha}\right)^{k-1}$$

$$=\frac{\alpha}{(1+\alpha)^2}\times 1\Big/\left[1-\frac{\alpha}{(1+\alpha)}\right]^2=\alpha.$$

同理,利用幂级数的求和技巧可得

$$E(X^2)=\sum_{k=1}^{\infty}k^2\frac{\alpha^k}{(1+\alpha)^{k+1}}=\frac{\alpha}{(1+\alpha)^2}\sum_{k=1}^{\infty}k^2\left(\frac{\alpha}{1+\alpha}\right)^{k-1}$$

$$=\frac{\alpha}{(1+\alpha)^2}\left(1+\frac{\alpha}{1+\alpha}\right)\Big/\left(1-\frac{\alpha}{1+\alpha}\right)^2=\alpha(1+2\alpha),$$

所以 $$D(X)=E(X^2)-[E(X)]^2=\alpha(1+\alpha).$$

例3 设 X_1,X_2,X_3 都服从$[0,2]$上的均匀分布,则 $E(3X_1-X_2+2X_3)=(\quad)$.

(A) 4; (B) 3; (C) 1; (D) 2.

解 选(A). 因为

$$f(x_i)=\begin{cases}1/2, & 0<x<2,\\ 0, & 其它,\end{cases}\quad E(X_i)=\int_0^2\frac{x}{2}\mathrm{d}x=1,$$

所以

$$E(3X_1-X_2+2X_3)=3E(X_1)-E(X_2)+2E(X_3)=4.$$

例4 设随机变量X的分布函数为

$$F(x)=\begin{cases}0, & x<-1,\\ a+b\arcsin x, & -1\leqslant x<1,\\ 1, & x\geqslant 1,\end{cases}$$

试确定常数a和b,并求$E(X)$,$D(X)$.

解 $f(x)=F'(x)$,所以

$$f(x)=\begin{cases}b/\sqrt{1-x^2}, & -1\leqslant x<1,\\ 0, & 其它,\end{cases}$$

$$\int_{-1}^{1}\frac{b}{\sqrt{1-x^2}}\mathrm{d}x=b\pi=1\Longrightarrow b=\frac{1}{\pi}.$$

由连续性,有

$$F(1)=a+\frac{1}{\pi}\arcsin x=a+\frac{1}{2}=1\Longrightarrow a=\frac{1}{2}.$$

所以 $$E(X)=\int_{-1}^{1}x\frac{1}{\pi\sqrt{1-x^2}}\mathrm{d}x\ (\text{由奇偶性})=0.$$

$$D(X)=E(X^2)=\int_{-1}^{1}x^2\frac{1}{\pi\sqrt{1-x^2}}\mathrm{d}x=\frac{2}{\pi}\int_{0}^{1}x^2\frac{1}{\sqrt{1-x^2}}\mathrm{d}x$$

$$=\frac{2}{\pi}\left(-\frac{x}{2}\sqrt{1-x^2}+\frac{1}{2}\arcsin x\right)\Big|_0^1=\frac{1}{2}.$$

例5 设随机变量 X 服从超几何分布.

$$P\{X=m\}=C_M^mC_{N-M}^{n-m}/C_N^n,\quad m=0,1,\cdots,n,$$

求 $E(X)$ 和 $D(X)$.

解 $$E(X)=\sum_{m=0}^{n}mC_M^mC_{N-M}^{n-m}/C_N^n.$$

由 $\sum\limits_{m=1}^{n}p_m=1$，故

$$\sum_{m=0}^{n}C_M^mC_{N-M}^{n-m}/C_N^n=1\ (n\leqslant N-M,n\leqslant M).$$

于是 $$\sum_{m=0}^{n}C_M^mC_{N-M}^{n-m}=C_N^n,\quad mC_M^m=MC_{M-1}^{m-1},$$

从而 $$\sum_{m=0}^{n}mC_M^mC_{N-M}^{n-m}=\sum_{m=1}^{n}MC_{M-1}^{m-1}C_{N-1-(M-1)}^{n-1-(m-1)}$$

$$=M\sum_{m=0}^{n-1}C_{M-1}^{m}C_{N-1-(m-1)}^{n-1-m}=MC_{N-1}^{n-1},$$

所以 $$E(X)=MC_{N-1}^{n-1}/C_N^n=Mn/N.$$

类似可求得

$$E(X^2)=\frac{n(n-1)M(M-1)}{N(N-1)}+\frac{nM}{N},$$

故 $$D(X)=E(X^2)-[E(X)]^2=\frac{nM(N-M)(N-n)}{N^2(N-1)}.$$

例6 设随机变量 $X\sim\pi(\lambda)$，且已知 $E[(X-1)(X-2)]=1$，则 λ $=$ ________.

解 由$X\sim\pi(\lambda)$知，$E(X)=\lambda, D(X)=\lambda$. 又

$$E[(X-1)(X-2)]=E[X^2-3X+2]=E(X^2)-3E(X)+2$$
$$=D(X)+[E(X)]^2-3E(X)+2$$
$$=\lambda+\lambda^2-3\lambda+2=1,$$

故
$$\lambda=1.$$

例7 设随机变量X,Y相互独立，且都服从$N(0,1/2)$，求随机变量$|X-Y|$的方差.

解 令$Z=X-Y$，则由参数可加性知，$Z\sim N(0,1)$，即

$$E(Z)=0,\quad D(Z)=1,\quad E(Z^2)=D(Z)=1.$$

由
$$D(|X-Y|)=D(|Z|)=E(|Z|^2)-[E(|Z|)]^2$$
$$=E(Z^2)-[E(|Z|)]^2,$$

$$E(|Z|)=\int_{-\infty}^{+\infty}|z|\frac{1}{\sqrt{2\pi}}e^{-z^2/2}dz=\frac{2}{\sqrt{2\pi}}\int_0^{+\infty}ze^{-z^2/2}dz=\sqrt{\frac{2}{\pi}}.$$

所以
$$D|X-Y|=1-2/\pi.$$

例8 设在某一规定的时间间隔内，一电气设备用于最大负荷的时间X(单位:min)是一个随机变量，概率密度是

$$f(x)=\begin{cases}x/1500^2, & 0\leqslant x\leqslant 1500,\\(3000-x)/1500^2, & 1500<x\leqslant 3000.\\0, & \text{其它},\end{cases}$$

求X的数学期望$E(X)$.

解 直接代入公式，分段计算积分，有

$$E(X)=\int_0^{1500}\frac{x^2}{1500^2}dx+\int_{1500}^{3000}\frac{x(3000-x)}{1500^2}dx$$
$$=\frac{x^3}{3\times1500^2}\Big|_0^{1500}+\frac{3000x^2/2-x^3/3}{1500^2}\Big|_{1500}^{3000}$$
$$=500+4500-3500=1500\ (\text{min}).$$

例9 设随机变量X的概率密度为

$$f(x)=\begin{cases}1-|1-x|, & 0<x<2,\\0, & \text{其它},\end{cases}$$

求 $E(X)$ 和 $D(X)$.

解 如图 4.1 所示，将绝对值记号去掉，化为 $f(x)=\begin{cases} x, & 0<x<1, \\ 2-x, & 1\leqslant x<2, \\ 0, & \text{其它}. \end{cases}$ 于是

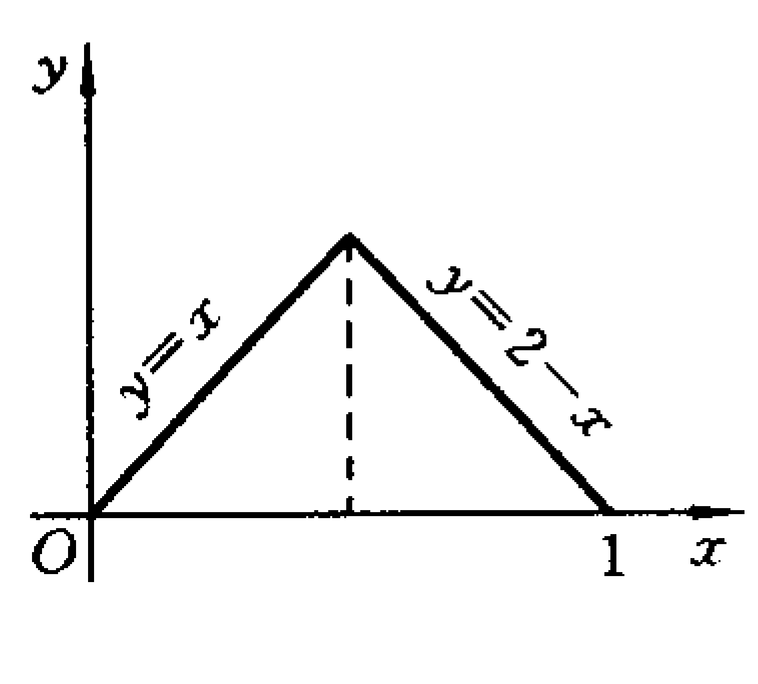


图 4.1

$$E(X)=\int_0^1 x\cdot x\mathrm{d}x+\int_1^2 x(2-x)\mathrm{d}x$$

$$=\frac{1}{3}+\frac{2}{3}=1,$$

$$E(X^2)=\int_0^1 x^2\cdot x\mathrm{d}x+\int_1^2 x^2(2-x)\mathrm{d}x=\frac{1}{4}+\frac{11}{12}=\frac{7}{6},$$

所以
$$D(X)=E(X^2)-[E(X)]^2=\frac{7}{6}-1=\frac{1}{6}.$$

例 10 设随机变量 X 服从拉普拉斯分布，概率密度为

$$f(x)=\frac{1}{2\lambda}\mathrm{e}^{-|x-\mu|/\lambda}\quad(\lambda>0),$$

求 $E(X)$ 和 $D(X)$.

解 $E(X)=\dfrac{1}{2\lambda}\displaystyle\int_{-\infty}^{+\infty}x\mathrm{e}^{-|x-\mu|/\lambda}\mathrm{d}x$ （令 $x-\mu=t$）

$$=\frac{1}{2\lambda}\int_{-\infty}^{+\infty}(t+\mu)\mathrm{e}^{-|t|/\lambda}\mathrm{d}t=\frac{1}{2\lambda}\int_{-\infty}^{+\infty}\mu\mathrm{e}^{-|t|/\lambda}\mathrm{d}t$$

$$=\frac{1}{\lambda}\int_0^{+\infty}\mu\mathrm{e}^{t/\lambda}\mathrm{d}t=\mu\ \left(\int_{-\infty}^{+\infty}t\mathrm{e}^{-|t|/\lambda}\mathrm{d}t=0\right),$$

$$E(X^2)=\frac{1}{2\lambda}\int_{-\infty}^{+\infty}x^2\mathrm{e}^{-|x-\mu|/\lambda}\mathrm{d}x=\frac{1}{2\lambda}\int_{-\infty}^{+\infty}(t+\mu)^2\mathrm{e}^{-|t|/\lambda}\mathrm{d}t$$

$$=\frac{1}{\lambda}\int_0^{+\infty}t\mathrm{e}^{-t/\lambda}\mathrm{d}t+\frac{\mu}{\lambda}\int_0^{\infty}\mathrm{e}^{-t/\lambda}\mathrm{d}t=2\lambda^2+\mu^2,$$

所以
$$D(X)=E(X^2)-[E(X)]^2=2\lambda^2.$$

例 11 设 X 服从贝塔(β)分布，概率密度为

$$f(x)=\begin{cases} \dfrac{1}{\mathrm{B}(p,q)}x^{p-1}(1-x)^{q-1}, & 0<x<1, \\ 0, & \text{其它}, \end{cases}$$

求 $E(X)$ 与 $D(X)$.

解 $B(p,q)=\int_0^1 x^{p-1}(1-x)^{q-1}\mathrm{d}x=\dfrac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)}$.

而 $\Gamma(\alpha)=\int_0^{\infty} x^{\alpha-1}\mathrm{e}^{-x}\mathrm{d}x$， $\Gamma(\alpha+1)=\alpha\Gamma(\alpha)$， $\Gamma(1)=1$，

所以

$$E(X)=\frac{1}{B(p,q)}\int_0^1 x^{p}(1-x)^{q-1}\mathrm{d}x=\frac{\Gamma(p+q)}{\Gamma(p)+\Gamma(q)}B(p+1,q)$$
$$=\frac{\Gamma(p+q)}{\Gamma(p)+\Gamma(q)}\cdot\frac{\Gamma(p+1)\Gamma(q)}{\Gamma(p+q+1)}$$
$$=\frac{\Gamma(p+q)}{\Gamma(p)+\Gamma(q)}\cdot\frac{p\Gamma(p)\Gamma(q)}{(p+q)\Gamma(p+q)}=\frac{p}{p+q}.$$

类似地

$$E(X^2)=\frac{1}{B(p,q)}\int_0^1 x^{p+1}(1-x)^{q}\mathrm{d}x=\frac{1}{B(p,q)}B(p+2,q)$$
$$=\frac{\Gamma(p+q)}{\Gamma(p)+\Gamma(q)}\cdot\frac{(p+1)p\Gamma(p)\Gamma(q)}{(p+q+1)(p+q)\Gamma(p+q)}$$
$$=\frac{(p+1)p}{(p+q+1)(p+q)},$$

所以 $D(X)=E(X^2)-[E(X)]^2=\dfrac{pq}{(p+q+1)(p+q)}$.

下面讨论二维随机变量与随机变量函数的情形.

例 12 设随机变量(X,Y)的联合分布律为

Y \ X	1	2
-1	1/4	1/2
1	0	1/4

求$E(X)$,$E(Y)$和$E(XY)$.

解 因为

X	1	2
p_k	1/4	3/4

Y	-1	1
p_k	3/4	1/4

所以

$$E(X)=1\times 1/4+2\times 3/4=7/4,$$
$$E(Y)=-1\times 3/4+1\times 1/4=-1/2,$$

$$
\begin{aligned}
E(XY)=&1\times(-1)\times1/4+1\times1\times0\\
&+2\times(-1)\times1/2+2\times1\times1/4\\
=&-3/4.
\end{aligned}
$$

例13 设随机变量 $X\sim N(0,1)$，$Y=2X+1$，则 $Y\sim$(　　).

(A) $N(0,1)$；　　(B) $N(1,1)$；

(C) $N(1,2)$；　　(D) $N(1,4)$.

解 选(D). 因为 $E(Y)=2E(X)+1=1$，$D(Y)=4D(X)=4$，所以 $Y\sim N(1,4)$.

例14 设随机变量 X 的概率密度为

$$f(x)=\frac{1}{2\sqrt{\pi}}\mathrm{e}^{-(x+3)^2/4},\quad -\infty<x<+\infty,$$

若 $Y\sim N(0,1)$，则 $Y=$(　　).

(A) $(X+3)/2$；　　(B) $(X+3)/\sqrt{2}$；

(C) $(X-3)/2$；　　(D) $(X-3)/\sqrt{2}$.

解 选(B). 因为 $X\sim N(-3,2)$，所以，当 $Y=\dfrac{X-\mu}{\sigma}=\dfrac{X+3}{\sqrt{2}}$ 时，$Y\sim N(0,1)$.

例15 某厂生产的一种设备的使用寿命 X 的概率密度为

$$f(x)=\begin{cases}\dfrac{1}{4}\mathrm{e}^{-x/4}, & x>0,\\ 0, & x\leqslant 0.\end{cases}$$

工厂规定：已售设备在一年内损坏予以调换. 若工厂出售一台设备盈利100元，调换一台设备厂方损失300元，求厂方出售一台设备盈利值的数学期望.

解 Y 可能值为100和-200(即$100-300$)，于是

$$P\{Y=100\}=P\{X\geqslant1\}=\int_1^{+\infty}\frac{1}{4}\mathrm{e}^{-x/4}\mathrm{d}x=\mathrm{e}^{-1/4},$$

$$P\{Y=-200\}=P\{X<1\}=1-\mathrm{e}^{1/4},$$

所以　　$E(Y)=100\mathrm{e}^{-1/4}+(-200)\times(1-\mathrm{e}^{-1/4})$

$$=300e^{-1/4}-200.$$

例 16 设某种出口商品的国际需求量 $X\sim U(2000,4000)$. 若每出售 1 t 商品可赚取外汇 3 万元，但积压 1 t 要花保养费 1 万元. 问：应组织多少货源，才能获取最大收益？

解 设商品量的最大值点为 x_0，收益为随机变量 Y，$Y=g(X)$，有

$$f(x)=\begin{cases}1/2000, & 2000<x<4000,\\ 0, & \text{其它},\end{cases}$$

$$Y=g(X)=\begin{cases}3x_0, & X>x_0,\\ 3X-(x_0-X), & X\leqslant x_0,\end{cases}$$

则
$$E(Y)=\int_{2000}^{x_0}(4x-x_0)\frac{1}{2000}\mathrm{d}x+\int_{x_0}^{4000}3x_0\frac{1}{2000}\mathrm{d}x$$
$$=\frac{1}{1000}(-x_0^2+7000x_0-4\times10^6).$$

因为，由 $[E(Y)]'_{x_0}=(-2x_0+7000)/1000=0$，得 $x_0=3500$；又 $[E(Y)]''_{x_0}=-2/1000$，所以 $x_0=3500$ 为最大值点. 从而知，组织货源 3500 t 时，收益最大.

例 17 设随机变量 $X\sim U(0,1)$，$Y\sim U(1,3)$，X 与 Y 相互独立，求 $E(XY)$ 与 $D(XY)$.

解 因为

$$f_X(x)=\begin{cases}1, & 0<x<1,\\ 0, & \text{其它},\end{cases}\qquad f_Y(y)=\begin{cases}1/2, & 1<y<3,\\ 0, & \text{其它},\end{cases}$$

所以
$$f(x,y)=\begin{cases}1/2, & 0<x<1,1<y<3,\\ 0, & \text{其它}.\end{cases}$$

于是
$$E(XY)=\int_0^1 x\mathrm{d}x\int_1^3\frac{1}{2}y\mathrm{d}y=\frac{1}{2}\times2=1,$$
$$E(X^2Y^2)=\int_0^1 x^2\mathrm{d}x\int_1^3\frac{1}{2}y^2\mathrm{d}y=\frac{1}{3}\times\frac{13}{3}=\frac{13}{9},$$

从而
$$D(XY)=E(X^2Y^2)-[E(XY)]^2=\frac{13}{9}-1=\frac{4}{9}.$$

例 18 在长为 l 的线段上任意选取两点，求：两点间距离 Z 的

数学期望与方差,并求 $E(Z^n)$ 和 $D(Z^n)$.

解 如图4.2所示,将线段置于区间 $[0,l]$ 上,则两点为随机变量 X,Y. X 与 Y 相互独立且同分布,服从 $U[0,l]$, $Z=|X-Y|$,因而

$$F(z)=P\{Z\leqslant z\}=P\{|X-Y|\leqslant z\}.$$

利用几何型概率可求得

$$\begin{aligned}&P\{|X-Y|\leqslant z\}\\&=P\{X-Z\leqslant Y\leqslant X+Z\}\\&=[l^2-(l-z)^2]/l^2\\&=1-(1-z/l)^2,\ 0<z<l.\end{aligned}$$

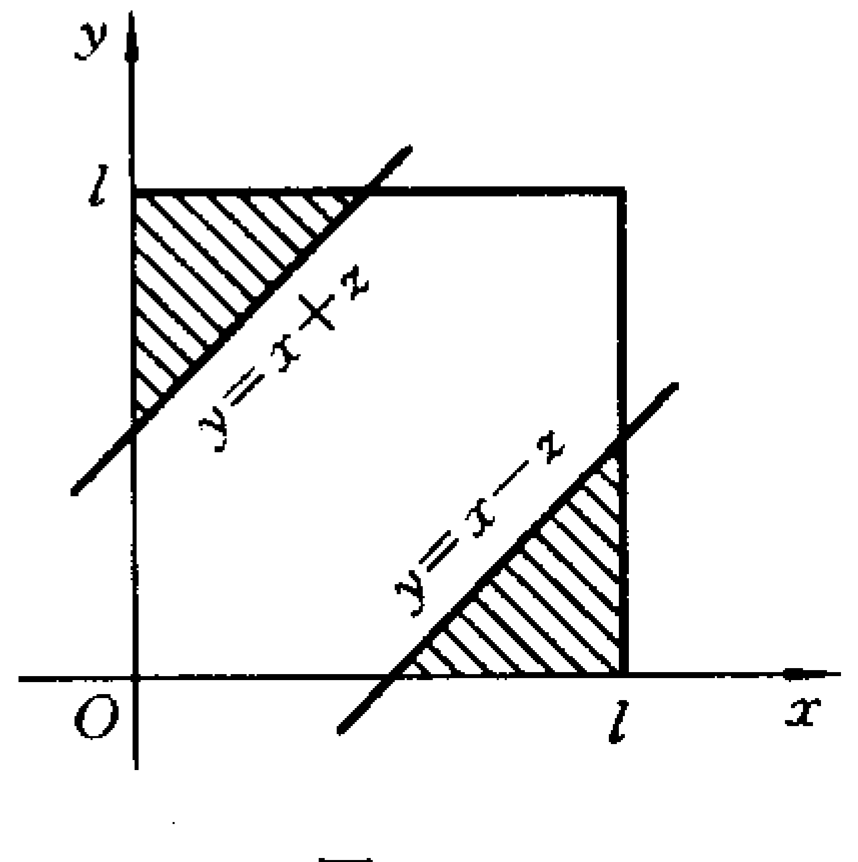


图 4.2

当 $z\leqslant 0$ 时, $F(z)=0$;当 $z\geqslant l$ 时, $F(z)=1$. 所以

$$f(z)=\begin{cases}2(1-z/l)/l, & 0<z<l,\\ 0, & \text{其它}.\end{cases}$$

于是 $$E(Z)=\frac{2}{l}\int_0^l z\left(1-\frac{z}{l}\right)\mathrm{d}z=\frac{2}{l}\left(\frac{z^2}{2}-\frac{z^3}{3l}\right)\bigg|_0^l=\frac{l}{3},$$

$$E(Z^2)=\frac{2}{l}\int_0^l z^2\left(1-\frac{z}{l}\right)\mathrm{d}z=\frac{l^2}{6},$$

$$E(Z^n)=\frac{2}{l}\int_0^l z^n\left(1-\frac{z}{l}\right)\mathrm{d}z=\frac{2l^n}{(n+1)(n+2)},$$

$$E(Z^{2n})=\frac{2l^{2n}}{(2n+1)(2n+2)},$$

所以 $$D(Z)=E(Z^2)-[E(Z)]^2=l^2/18,$$

$$\begin{aligned}D(Z^n)&=E(Z^{2n})-[E(Z^n)]^2\\&=\left[\frac{2}{(2n+1)(2n+2)}-\frac{4}{(n+1)^2(n+2)^2}\right]l^{2n}.\end{aligned}$$

例 19 设随机变量 $X\sim N(0,\sigma^2)$,求 $E(X^n)$.

解 先进行标准化代换,令 $Y=\dfrac{X}{\sigma}$,则 $Y\sim N(0,1)$, $f(y)=\dfrac{1}{\sqrt{2\pi}}\mathrm{e}^{-y^2/2}$, $X^n=\sigma^nY^n$. 于是

$$E(X^n)=\sigma^n E(Y^n)=\frac{\sigma^n}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} y^n \mathrm{e}^{-y^2/2}\mathrm{d}y.$$

当n为奇数时，$E(X^n)=0$；当n为偶数时，令$u=y^2/2$，于是

$$\begin{aligned}E(X^n)&=\sigma^n 2^{n/2}\Big/\sqrt{2\pi}\cdot\int_0^{+\infty} u^{(n+1)/2-1}\mathrm{e}^{-u}\mathrm{d}y\\&=\sigma^n 2^{n/2}\Big/\sqrt{2\pi}\cdot\Gamma\left(\frac{n+1}{2}\right)\\&=\sigma^n 2^{n/2}\Big/\sqrt{2\pi}\cdot\left(\frac{n-1}{2}\right)\left(\frac{n-3}{2}\right)\cdots\left(\frac{1}{2}\right)\Gamma\left(\frac{1}{2}\right)\\&=\sigma^n(n-1)!!,\end{aligned}$$

所以 $$E(X^n)=\begin{cases}\sigma^n(n-1)!!, & n\text{ 为大于 2 的偶数},\\ 0, & n\text{ 为奇数}.\end{cases}$$

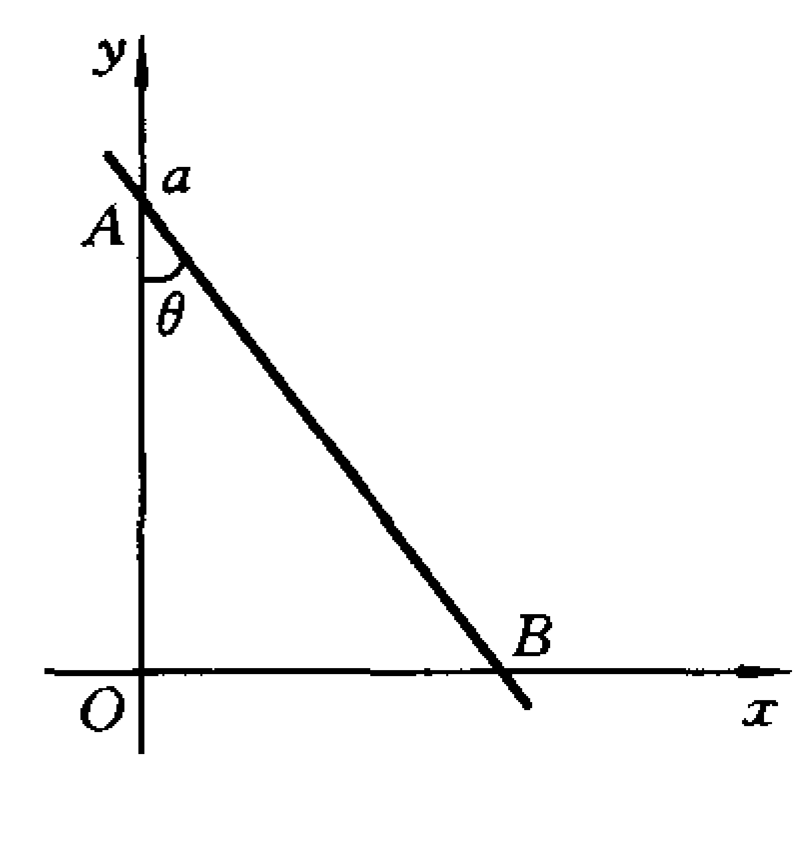


图 4.3

例 20 设平面上点A的坐标为$(0,a)$，$a>0$. 过点A的直线l与y轴的夹角为θ，l交x轴于点B(见图 4.3). 已知Θ在$[0,\pi/4]$上均匀分布，求$\triangle OAB$面积的数学期望.

解 设随机变量Θ的概率密度是

$$f(\theta)=\begin{cases}4/\pi, & 0\leqslant\theta\leqslant\pi/4,\\ 0, & \text{其它},\end{cases}$$

而面积$S=\frac{a}{2}\cdot a\tan\theta$，是随机变量$\Theta$的函数，故

$$E(S)=\frac{a^2}{2}\int_0^{\pi/4}\tan\theta\cdot\frac{4}{\pi}\mathrm{d}\theta=\frac{2a^2}{\pi}(-\ln\cos\theta)\Big|_0^{\pi/4}=\frac{a^2}{\pi}\ln 2.$$

例 21 随机变量X与Y相互独立，且都服从正态分布$N(a,\sigma^2)$，求$E[\max(X,Y)]$与$E[\min(X,Y)]$.

解 (1) 令$U=\frac{X-a}{\sigma}$，$V=\frac{Y-a}{\sigma}$，则U与V相互独立，且都服从$N(0,1)$.

$$\max(X,Y)=a+\sigma\max(U,V),$$

$$f_{UV}(u,v)=\frac{1}{2\pi}\mathrm{e}^{-(u^2+v^2)/2}.$$

所以

$$E[\max(U,V)]=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\max(u,v)\cdot\frac{1}{2\pi}\mathrm{e}^{-(u^2+v^2)/2}\mathrm{d}u\mathrm{d}v$$

$$=\iint_{u-v<0}vf(u,v)\mathrm{d}u\mathrm{d}v+\iint_{u-v\geqslant 0}uf(u,v)\mathrm{d}u\mathrm{d}v$$

$$=\frac{1}{2\pi}\left(\int_{-\infty}^{+\infty}\mathrm{e}^{-u^2/2}\mathrm{d}u\int_{u}^{+\infty}v\mathrm{e}^{-v^2/2}\mathrm{d}v\right.$$

$$\left.+\int_{-\infty}^{+\infty}\mathrm{e}^{-v^2/2}\mathrm{d}v\int_{v}^{+\infty}u\mathrm{e}^{-u^2/2}\mathrm{d}u\right)$$

$$=\frac{1}{\pi}\int_{-\infty}^{+\infty}\mathrm{e}^{-u^2}\mathrm{d}u=\frac{1}{\sqrt{\pi}}.$$

故 $$E[\max(X,Y)]=a+\sigma\frac{1}{\sqrt{\pi}}.$$

(2) 类似地,有

$$\min(X,Y)=a+\sigma\min(U,V),$$

$$E[\min(U,V)]=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\min(u,v)\cdot\frac{1}{2\pi}\mathrm{e}^{-(u^2+v^2)/2}\mathrm{d}u\mathrm{d}v$$

$$=\iint_{u-v<0}uf(u,v)\mathrm{d}u\mathrm{d}v+\iint_{u-v\geqslant 0}vf(u,v)\mathrm{d}u\mathrm{d}v$$

$$=\frac{1}{2\pi}\left[\int_{-\infty}^{+\infty}\mathrm{e}^{-v^2/2}\mathrm{d}v\int_{-\infty}^{v}u\mathrm{e}^{-u^2/2}\mathrm{d}u\right.$$

$$\left.+\int_{-\infty}^{+\infty}\mathrm{e}^{-u^2/2}\mathrm{d}u\int_{-\infty}^{u}v\mathrm{e}^{-v^2/2}\mathrm{d}v\right]$$

$$=\frac{-1}{\pi}\int_{-\infty}^{+\infty}\mathrm{e}^{-v^2}\mathrm{d}v=-\frac{1}{\sqrt{\pi}},$$

故 $$E[\min(X,Y)]=a-\sigma\frac{1}{\sqrt{\pi}}.$$

例22 从 $1,2,\cdots,N$ 中依次(不重复)取两个数,分别记为 X 和 Y,求 $E(X+Y)$.

解 显然,X 和 Y 相互独立且同分布,于是

$$E(X+Y)=2E(X)=2\left(1\times\frac{1}{N}+2\times\frac{1}{N}+\cdots+N\times\frac{1}{N}\right)$$

$$=\frac{2}{N}\times\frac{N(N+1)}{2}=N+1.$$

例 23 设随机变量 (X,Y) 服从 $G=\{(x,y)|\{y\geqslant 0,x^2+y^2\leqslant 1\}\}$ 上的均匀分布. 定义随机变量 U,V 如下：

$$U=\begin{cases}0, & X<0,\\ 1, & 0\leqslant X<Y,\\ 2, & X\geqslant Y,\end{cases}\qquad V=\begin{cases}0, & X\geqslant\sqrt{3}Y,\\ 1, & X<\sqrt{3}Y,\end{cases}$$

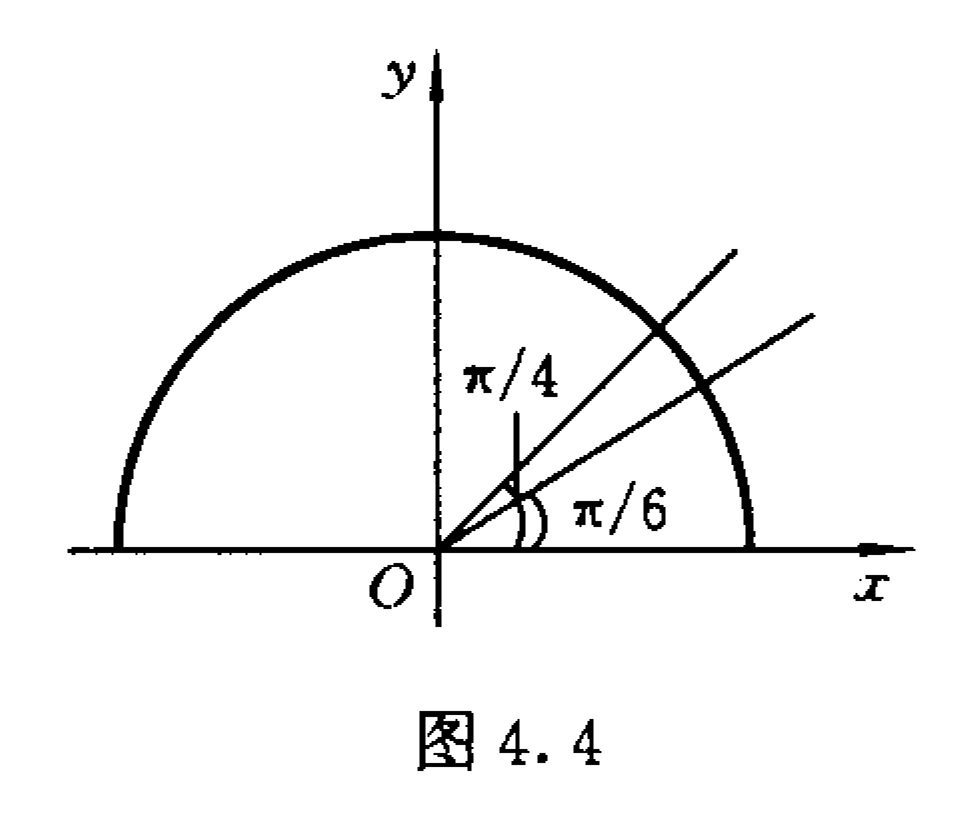


图 4.4

求 (U,V) 的联合概率分布及 $E(UV)$，$P\{UV=0\}$.

解 由图 4.4 知

$$f(x,y)=\begin{cases}2/\pi, & (x,y)\in G,\\ 0, & \text{其它},\end{cases}$$

得
$$P\{U=0,V=0\}=0,$$
$$P\{U=1,V=0\}=0,$$
$$P\{U=1,V=1\}=P\{0\leqslant X<Y\}=1/4,$$
$$P\{U=0,V=1\}=P\{X\leqslant 0\}=1/2,$$
$$P\{U=2,V=0\}=P\{X\geqslant\sqrt{3}Y\}=1/6,$$
$$P\{U=2,V=1\}=P\{Y\leqslant X<\sqrt{3}Y\}=1/4-1/6=1/12.$$

所以 (U,V) 的联合分布律为

V \ U	0	1	2
0	0	0	1/6
1	1/2	1/4	1/12

于是
$$E(UV)=1\times1\times1/4+2\times2\times1/12=5/12,$$
$$P\{UV=0\}=1/6+1/2=2/3.$$

例 24 设随机变量 $X\sim N(\mu,\sigma^2)$，$Y=\mathrm{e}^X$ 称为服从对数正态分布，求 $E(Y)$ 与 $D(Y)$.

解 可以由 Y 的密度函数

$$f(y)=\begin{cases}\dfrac{1}{\sigma y\sqrt{2\pi}}\mathrm{e}^{-(\ln y-\mu)^2/(2\sigma^2)}, & y>0,\\ 0, & \text{其它}\end{cases}$$

直接求解，也可以由 X 的概率密度间接求解，即

$$E(Y)=E(e^X)=\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{+\infty}e^x\cdot e^{-(x-\mu)^2/(2\sigma^2)}dx \quad \left(令\frac{x-\mu}{\sigma}=u\right)$$

$$=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}e^{\sigma u+\mu}\cdot e^{-u^2/2}du=e^{\mu+\sigma^2/2}\int_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi}}e^{-(u-\mu)^2/2}du$$

$$=e^{\mu+\sigma^2/2},$$

$$E(Y^2)=\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{+\infty}e^{2x}\cdot e^{-(x-\mu)^2/(2\sigma^2)}dx=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}e^{2(\sigma u+\mu)}\cdot e^{-u^2/2}du$$

$$=e^{2(\mu+\sigma^2)}\int_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi}}e^{-(u-2\sigma)^2/2}du=e^{2(\mu+\sigma^2)},$$

所以

$$D(Y)=E(Y^2)-[E(Y)]^2=e^{2(\mu+\sigma^2)}-e^{2\mu+\sigma^2}=e^{2\mu+\sigma^2}(e^{\sigma^2}-1).$$

例25 掷两枚骰子，以X 记第一枚骰子掷出的点数，以Y 记第二枚骰子掷出的点数，求 $E(X+Y)$和 $E(XY)$.

解 X,Y 相互独立且同分布，

$$P\{X=k\}=1/6,\ k=1,2,\cdots,6,$$

所以

$$E(X)=E(Y)=\frac{1}{6}(1+2+3+4+5+6)=\frac{7}{2},$$

$$E(X+Y)=2E(X)=7,$$

$$E(XY)=E(X)E(Y)=49/4=12.25.$$

例26 卡车装运水泥，设每袋水泥的重量(单位:kg)X 服从 $N(50,2.5^2)$，问:最多装多少袋水泥使总重量超过2000 kg 的概率大于0.05?

解 以 X_i 表示第 i 袋水泥的重量，X_i 相互独立且同分布，$X_i\sim N(50,2.5^2)$，$i=1,2,\cdots,n$，则 n 袋水泥的总重量 $Y=\sum_{i=1}^{n}X_i\sim N(50n,n(2.5)^2)$，问题化为求 n.

$$P\{X>2000\}=1-P\{X\leqslant 2000\}=1-\Phi\left(\frac{2000-50n}{2.5\sqrt{n}}\right)\leqslant 0.05,$$

即要$\frac{2000-50n}{2.5\sqrt{n}}\leqslant 1.645$，解得$n\leqslant 39.48$，故取$n=39$.

这里，容易犯的错误是$Y\sim N(n\mu,n^2\sigma^2)$，这时求出$n=36$，是不对的.

例27 设$X_1,X_2,\cdots,X_n$相互独立且同分布，$P\{X_i=0\}=1-p=q$，$P\{X_i=1\}=p$ $(0<p<1)$，以X记连续出现"1"或连续出现"0"的个数(称为游程)，求$E(X)$.

解 X的可取值为$1,2,\cdots,n$，则

$$P\{X=k\}=p_k=p^kq+q^kp,\quad k=1,2,\cdots,q.$$

于是
$$E(X)=\sum_{k=1}^{\infty}kp_k=\sum_{k=1}^{\infty}kp^kq+\sum_{k=1}^{\infty}kq^kp.$$

由级数知识知

$$\sum_{k=1}^{\infty}p_k=pq\Big(\sum_{k=1}^{\infty}p^{k-1}+\sum_{k=1}^{\infty}q^{k-1}\Big)=\frac{pq}{1-p}+\frac{pq}{1-q}=1,$$

$$\sum_{k=1}^{\infty}kp^{k-1}=\Big(\sum_{k=1}^{\infty}p^k\Big)'_k=\Big(\frac{p}{1-p}\Big)'_k=\frac{1}{(1-p)^2},$$

所以
$$E(X)=pq\Big(\sum_{k=1}^{\infty}kp^{k-1}+\sum_{k=1}^{\infty}kq^{k-1}\Big)$$

$$=\frac{pq}{(1-p)^2}+\frac{pq}{(1-q)^2}=\frac{p}{1-p}+\frac{1-p}{p}.$$

例28 某商场经销的一种商品的进货量X与顾客的需求量Y是相互独立的随机变量，都服从区间[10,20]上的均匀分布.设商店每售出一单位商品可获利润1000元，若需求量超过进货量，商店可到其它商店调剂，此时每单位商品可获利润500元；若进货量超过需求量，商场需削价处理，此时每单位商品亏损200元.求商场经营该种商品获利的期望值.

解 以Z记经营该种商品的利润值，则

$$Z=\begin{cases}1000Y-200(X-Y)=1200Y-200X, & Y\leqslant X,\\ 1000X+500(Y-X)=500(X+Y), & Y>X.\end{cases}$$

而X,Y的联合概率密度是

$$f(x,y)=\begin{cases}1/100, & 10\leqslant x,y\leqslant 20,\\ 0, & \text{其它}.\end{cases}$$

所以

$$E(Z)=\int_{10}^{20}\mathrm{d}x\int_{10}^{x}(12y-2x)\mathrm{d}y+\int_{10}^{20}\mathrm{d}y\int_{10}^{y}5(x+y)\mathrm{d}x$$

$$=\int_{10}^{20}(4x^2+20x-600)\mathrm{d}x+\int_{10}^{20}5\left(\frac{3}{2}y^2-10y-50\right)\mathrm{d}y$$

$$=19000/3+7500=13833.3\ (\text{元}).$$

例 29 设在 n 次独立重复试验中，每次试验的成功率为 p. 又设随机变量 Y 当成功偶数次时取 0，成功奇数次时取 1，求 $E(Y)$ 和 $D(Y)$.

解 以 X 记 n 次独立重复试验中成功的次数，则

$$P\{X=k\}=C_n^k p^k q^{n-k},\quad k=0,1,\cdots,n.$$

$$P\{Y=1\}=\frac{1}{2}\left\{\sum_{k=0}^{n}[P\{X=k\}+(-1)^k P\{X=k\}]\right\}$$

$$=\frac{1}{2}\left[\sum_{k=1}^{n}C_n^k p^k q^{n-k}+\sum_{k=1}^{n}C_n^k(-p)^k q^{n+k}\right]$$

$$=\frac{1}{2}[(p+q)^n+(p-q)^n]=\frac{1}{2}[1+(1-2p)^n].$$

从而

$$E(Y)=0\times P\{Y=0\}+1\times P\{Y=1\}=\frac{1}{2}[1+(1-2p)^n],$$

$$E(Y^2)=0\times P\{Y=0\}+1\times P\{Y=1\}=\frac{1}{2}[1+(1-2p)^n],$$

所以 $$D(Y)=E(Y^2)-[E(Y)]^2=\frac{1}{4}[1-(1-2p)^{2n}].$$

例 30 设某狩猎区内有 n 只狐狸，一共 r 次设若干个陷阱猎狐. 若在每次猎取中，对每只尚未捕获的狐狸而言，它落入陷阱的概率都是 p $(0<p<1)$，求第 r 次设陷阱捕获狐狸的数学期望.

解 令 $$X_i=\begin{cases}1, & \text{第 } i \text{ 只狐狸在第 } r \text{ 次捕获},\\ 0, & \text{其它},\end{cases}$$

则 $$E(X_i)=P\{X_i=1\}=(1-p)^{r-1}p$$

(前 $r-1$ 次未捕获，第 r 次被捕获)，

$$E(Y)=E\left(\sum_{i=1}^{n}X_i\right)=nE(X_i)=n(1-p)^{r-1}p.$$

二、分布未知时，求数学期望与方差

例 31　一个有 n 把钥匙的人要开他的门，他随机而独立地试开. 若其中只有一把钥匙能开门，试求：

(1) 把试开不成功的钥匙立即除去情形试开次数的数学期望与方差；

(2) 把试开不成功的钥匙不除去情形试开次数的数学期望与方差.

解　(1) 以 X 记试开的次数，可取值 $1,2,\cdots,n$. $\{X=i\}$ 表示前 $i-1$ 次没打开，第 i 次才打开事件，所以

$$P\{X=i\}=\frac{n-1}{n}\cdot\frac{n-2}{n-1}\cdot\cdots\cdot\frac{n-i+1}{n-i+2}\cdot\frac{1}{n-i+1}=\frac{1}{n},$$

$$E(X)=\sum_{i=1}^{n}i\cdot\frac{1}{n}=\frac{1}{n}\cdot\frac{n(n+1)}{2}=\frac{n+1}{2},$$

$$E(X^2)=\sum_{i=1}^{n}i^2\frac{1}{n}=\frac{1}{n}\cdot\frac{n(n+1)(2n+1)}{6}=\frac{(n+1)(2n+1)}{6},$$

故
$$D(X)=E(X^2)-[E(X)]^2=\frac{n^2-1}{12}.$$

(2) 以 Y 记试开的次数，则 Y 可取值 $1,2,\cdots,n,\cdots$. 与 X 可取值不同，$\{Y=i\}$ 表示前 $i-1$ 次没打开，第 i 次才打开事件. 所以

$$P\{Y=i\}=\left(\frac{n-1}{n}\right)^{i-1}\frac{1}{n}=\frac{1}{n}\left(1-\frac{1}{n}\right)^{i-1},$$

$$E(X)=\sum_{i=1}^{\infty}\frac{i}{n}\left(1-\frac{1}{n}\right)^{i-1}=\frac{1}{n}\sum_{i=1}^{\infty}p\left(1-\frac{1}{n}\right)^{i-1},$$

利用级数求和技巧(令 $(1-1/n)=p$，如例 27)，可得 $E(X)=n$.

类似地，
$$E(X^2)=2n^2-n.$$

故
$$D(E)=E(X^2)-[E(X)]^2=n(n-1).$$

例32　某射手对一目标进行射击，直到击中为止. 设每次射击的命中率为 p，求该射手射击次数的数学期望与方差.

解 显然，以 X 记射击的次数时，$P\{X=k\}=p(1-p)^{k-1}$（是几何分布，$0<p<1$）. 令 $1-p=q$，有

$$E(X)=\sum_{k=1}^{\infty}kpq^{k-1}=\frac{p}{(1-q)^2}=\frac{1}{p}\text{（利用级数求和技巧）}.$$

类似地，$$E(X^2)=\sum_{k=1}^{n}k^2pq^{k-1}=p\frac{1+q}{(1-q)^3}=\frac{2-p}{p^2},$$

所以 $$D(X)=E(X^2)-[E(X)]^2=\frac{2-p}{p^2}-\frac{1}{p^2}=\frac{1-p}{p^2}.$$

例 33 有 100 名战士参加实弹练习，设每名战士一次射击的命中率为 0.8，规定每名战士至多射击 4 次，若已射中则不再射击. 问：该次练习，至少应准备多少发子弹？

解 以 X_i 表示第 i 名战士需用的子弹数，$i=1,2,\cdots,100$. X_i 可取值为 1，2，3，4，则所求为 $E(X)=E\left(\sum_{i=1}^{100}X_i\right)=\sum_{i=1}^{100}E(X_i)$. 因为 X_i 相互独立且同分布，所以

$$P\{X_i=1\}=0.8,\quad P\{X_i=2\}=0.2\times0.8=0.16,$$
$$P\{X_i=3\}=0.2^2\times0.8=0.032,$$
$$P\{X_i=4\}=0.2^3\times0.8=0.008,$$

故 $$E(X_i)=0.8+2\times0.16+3\times0.032+4\times0.008=1.248.$$

于是 $$E(X)=\sum_{i=1}^{100}E(X_i)=100\times1.248=124.8,$$

即至少应准备 125 发子弹.

例34 一箱中有10 只同型的配件，其中2 只是次品. 装配工在使用时任取 1 只，若是废品，则扔掉重取 1 只，直到取得正品为止. 求在取得正品前，已取得次品数的数学期望.

解 以 X 记取得正品前已取得的次品数，则 X 可取值为 0，1，2，且

$$P\{X=0\}=\frac{8}{10},\quad P\{X=1\}=\frac{2}{10}\times\frac{8}{9}=\frac{8}{45},$$

$$P\{X=2\}=\frac{2}{10}\times\frac{1}{9}\times1=\frac{1}{45},$$

所以
$$E(X)=0+1\times\frac{8}{45}+2\times\frac{1}{45}=\frac{2}{9},$$
$$E(X^2)=0+1\times\frac{8}{45}+4\times\frac{1}{45}=\frac{4}{15},$$
$$D(X)=E(X^2)-[E(X)]^2=\frac{4}{15}-\left(\frac{2}{9}\right)^2=\frac{88}{405}.$$

例 35 将 3 个球独立地随机地放入 4 个编号为 1,2,3,4 的盒子中去. 以 X 表示其中至少有一个球的盒子的最小号码,求 $E(X)$.

解 先用古典型概率方法求 X 的分布律. 因为样本空间的基本事件数为 4^3,$X=1$ 含事件数为 4^3-3^3,$X=2$ 含事件数为 3^3-2^3,$X=3$ 含事件数为 2^3-1,$X=4$ 含事件数为 1,所以
$$P\{X=1\}=\frac{37}{64},\quad P\{X=2\}=\frac{19}{64},$$
$$P\{X=3\}=\frac{7}{64},\quad P\{X=4\}=\frac{1}{64}.$$
于是
$$E(X)=\frac{37}{64}+2\times\frac{19}{64}+3\times\frac{7}{64}+4\times\frac{1}{64}=\frac{100}{64}=\frac{25}{16}.$$

这里 $X=k$ 的事件数是这样算出来的:$X=1$ 含的事件数是将 3 个球任意放入 4 个盒子的事件数 4^3 减去任意放入除 1 号盒子的 3 个盒子的事件数 3^3.

例 36 设袋中装有编有号码为 $1,2,\cdots,n$ 的球,第 k 号的有 k 个. 现从中摸出一球,求所出现号码的数学期望.

解 以 X 记所摸得球的号码,则
$$P\{X=k\}=\frac{k}{1+2+\cdots+n}=\frac{2k}{n(n+1)},\quad k=1,2,\cdots,n,$$
故
$$E(X)=\sum_{k=1}^{n}kP\{X=k\}=\frac{2}{n(n+1)}\sum_{k-1}^{n}k^2$$
$$=\frac{2}{n(n+1)}\frac{n(n+1)(2n+1)}{6}=\frac{2n+1}{3}.$$

例 37 某地有 A、B 两队进行乒乓球比赛,规定一方先胜三盘,则比赛结束. 设每场比赛 A 队获胜的概率 $p=1/2$,以 X 记比赛

的盘数，求$E(X)$.

解 因为A、B两队的胜率相等，所以只需讨论A队获胜的情形就可以了. X的可能取值为3,4,5，求出X的分布律.

$$P\{X=3\}=2\times p^3=1/4,$$

$$P\{X=4\}=2pC_3^2p^2(1-p)=2\times 3/16=3/8,$$

$$P\{X=5\}=2pC_4^2p^2(1-p)^2=2\times 6\times 1/32=3/8,$$

所以 $$E(X)=3\times 1/4+4\times 3/8+5\times 3/8=33/8\ (\text{盘}).$$

例38 某产品的次品率为0.1，检验员每天检验4次，每次随机地取10件产品进行检验. 如发现其中的次品数多于1，就去调整设备. 以X表示1天中调整设备的次数，求$E(X)$.

解一 一般的方法是先求设备需要调整的概率. 若以Y记每次检验中发现的次品件数，则$Y\sim B(10,0.1)$. 当$Y>1$时，需调整设备. 所以

$$\begin{aligned}P\{Y>1\}&=1-P\{Y\leqslant 1\}\\&=1-(1-0.1)^{10}-C_{10}^1\times 0.1\times(1-0.1)^9\\&=1-0.9^{10}-0.9^9=0.264.\end{aligned}$$

于是，$X\sim B(4,0.264)$，从而

$$E(X)=4\times 0.264=1.056\ (\text{次}).$$

解二 技巧性较强的方法是利用数学期望的性质求解. 设

$$X_i=\begin{cases}0, & \text{第}\,i\,\text{次检验发现次品件数}\leqslant 1,\\1, & \text{第}\,i\,\text{次检验发现次品件数}>1,\end{cases}$$

则$X=\sum\limits_{i=1}^{4}X_i$，而

$$P\{X_i=1\}=1-0.9^{10}-C_{10}^1\times 0.1\times 0.9^9=0.264,$$

$$E(X)=0\times P\{X_i=0\}+1\times P\{X_i=1\}=0.264,$$

所以 $$E(X)=\sum_{i=1}^{4}E(X_i)=4\times 0.264=1.056\ (\text{次}).$$

这种方法经常用来求较复杂的但能分解为若干较简单的同分布的随机变量和的随机变量的数字特征.

例 39 某机场的送客车一次载 20 名旅客自机场开出，沿途有 10 个停车点，若到达停车点无人下车则车不停. 设每名旅客在各个停车点下车是等可能的，求送客车停车次数的数学期望.

解 利用上题解二的方法. 设

$$X_i=\begin{cases}0, & \text{第}\,i\,\text{个停车点无人下车},\\ 1, & \text{第}\,i\,\text{个停车点有人下车},\end{cases}$$

$i=1,2,\cdots,10$. 送客车停车的总次数 $X=\sum\limits_{i=1}^{10}X_i$，则

$$P\{X_i=0\}=P\{\text{第}\,i\,\text{个停车点无人下车}\}=\left(\frac{9}{10}\right)^{20},$$

$$P\{X_i=1\}=P\{\text{第}\,i\,\text{个停车点至少有一个下车}\}=1-\left(\frac{9}{10}\right)^{20}.$$

所以 $E(X_i)=0\times\left(\frac{9}{10}\right)^{20}+1\times\left[1-\left(\frac{9}{10}\right)^{20}\right]=1-\left(\frac{9}{10}\right)^{20}$，

$$E(X)=E\left(\sum_{i=1}^{10}X_i\right)=\sum_{i=1}^{10}E(X_i)=10\left[1-\left(\frac{9}{10}\right)^{20}\right]=8.784\ (\text{次}).$$

需要提醒的是，在用该方法解题时，各 X_i 应当是相互独立的，其独立性应易于确定.

例 40 一袋中装有 3 个红球、5 个白球，从中抽取 4 次，每次抽一球，以 X 记取到红球的次数. 求：

(1) 在放回抽样下，X 的分布律为 $P\{X=k\}$，$E(X)$ 及 X 的最可能取值；

(2) 在无放回抽样下，X 的分布律为 $P\{X=k\}$，$E(X)$.

解 (1) 在放回抽样下，每次取到红球的概率为 3/8，所以 $X\sim B(4,3/8)$. 因此

$$P\{X=k\}=C_4^k\times(3/8)^k\times(5/8)^{4-k},\quad k=0,1,\cdots,4,$$

$$E(X)=np=4\times 3/8=3/2,$$

X 的最可能取值为 $[np+p]=[3/2+3/8]=[15/8]=1$.

(2) 在无放回抽样下，抽得红球的概率服从超几何分布 $H(4,3,8)$，所以

$$P\{X=k\}=C_3^k C_5^{4-k}/C_8^4,\quad k=0,1,2,3,$$
$$E(X)=nM/N=4\times 3/8=3/2.$$

例41 某城镇有N辆汽车，车牌号的尾数从1到N. 现在在街上随机地记下n辆车牌号的尾数，设其最大号码为X，求$E(X)$（只考虑放回抽样情形）.

解 先求出X的概率分布. 因为基本事件总数为N^n，有利事件数为$k^n-(k-1)^n$（$X=k$时，号码不大于k的有k^n种，号码不大于$(k-1)$的有$(k-1)^n$种），所以
$$P\{X=k\}=[k^n-(k-1)^n]/N^n,$$
于是
$$E(X)=\sum_{k=1}^{N}kP\{X=k\}=\frac{1}{N^n}\sum_{k=1}^{N}k[k^n-(k-1)^n]$$
$$=\frac{1}{N^n}\{1(1^n-0^n)+2(2^n-1^n)+\cdots+N[N^n-(N-1)^n)]\}$$
$$=\frac{-1}{N^n}[1^n+2^n+\cdots+(N-1)^n-N^{n+1}]$$
$$=N-\frac{1}{N^n}\sum_{k=1}^{N-1}k^n.$$

例42 已知随机变量X的概率密度形式为
$$f(x)=\begin{cases}ax, & 0<x<2,\\ cx+b, & 2\leqslant x\leqslant 4,\\ 0, & \text{其它},\end{cases}$$
且$E(X)=2$，$P\{1<X<3\}=3/4$，求：

(1) a,b,c的值； (2) $E(Y)=E(e^X)$.

解 (1) 由$1=\int_{-\infty}^{+\infty}f(x)dx=\int_0^2 ax dx+\int_2^4(cx+b)dx$，

得 $$2a+6c+2b=1.$$

由 $$2=\int_0^2 xax dx+\int_2^4 x(cx+b)dx,$$

得 $$8a/3+56c/3+6b=2.$$

由 $$3/4=P\{1<X<3\}=\int_1^2 ax\mathrm{d}x+\int_2^3(cx+b)\mathrm{d}x,$$

得 $$3/2+5c/2+b=3/4.$$

解方程组

$$\begin{cases}2a+6c+2b=1,\\ 8a/3+56c/3+6b=2,\\ 3/2+5c/2+b=3/4,\end{cases}\quad 得\quad \begin{cases}a=1/4,\\ b=1,\\ c=-1/4.\end{cases}$$

(2) $$E(Y)=E(\mathrm{e}^X)=\int_0^2\mathrm{e}^x\cdot\frac{x}{4}\mathrm{d}x+\int_2^4\mathrm{e}^x\left(1-\frac{x}{4}\right)\mathrm{d}x$$

$$=\frac{1}{4}(\mathrm{e}^2-1)^2.$$

例 43 设 $X_1,X_2,\cdots,X_n$ 是相互独立的随机变量，且有

$$E(X_i)=\mu,\quad D(X_i)=\sigma^2,\quad i=1,2,\cdots,n.$$

记 $\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$，$S^2=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2$，验证：

(1) $E(\overline{X})=\mu$，$D(\overline{X})=\sigma^2/n$； (2) $S^2=\frac{1}{n-1}\left(\sum X_i^2-n\overline{X}^2\right)$；

(3) $E(S^2)=\sigma^2$.

解 (1) $$E(\overline{X})=E\left(\frac{1}{n}\sum_{i=1}^{n}X_i\right)=\frac{1}{n}\sum_{i=1}^{n}E(X_i)=\frac{1}{n}n\mu=\mu,$$

$$D(\overline{X})=D\left(\frac{1}{n}\sum_{i=1}^{n}X_i\right)=\frac{1}{n^2}\sum_{i=1}^{n}D(X_i)=\frac{1}{n^2}n\sigma^2=\frac{1}{n}\sigma^2.$$

(2) $$S^2=\frac{1}{n-1}\sum_{i=1}^{n}(X_i^2-2X_i\overline{X}+\overline{X}^2)$$

$$=\frac{1}{n-1}\left[\sum_{i=1}^{n}X_i^2-\sum_{i=1}^{n}\overline{X}(2X_i-\overline{X})\right]$$

$$=\frac{1}{n-1}\left[\sum_{i=1}^{n}X_i^2-\overline{X}\left(2\sum_{i=1}^{n}X_i-n\overline{X}\right)\right]$$

$$=\frac{1}{n-1}\left(\sum_{i=1}^{n}X_i^2-n\overline{X}^2\right).$$

(3) $$E(S^2)=E\left[\frac{1}{n-1}\left(\sum_{i=1}^{n}X_i^2-n\overline{X}^2\right)\right]$$

$$=\frac{1}{n-1}E\left\{\sum_{i=1}^{n}[(X_i-\mu)-(\overline{X}-\mu)]^2\right\}$$

$$=\frac{1}{n-1}\left\{\sum_{i=1}^{n}E(X_i-\mu)^2-nE[(\overline{X}-\mu)^2]\right\}$$

$$=\frac{1}{n-1}\left(n\sigma^2-n\cdot\frac{\sigma^2}{n}\right)=\frac{1}{n-1}(n-1)\sigma^2=\sigma^2.$$

例 44 设随机变量 X 满足

$$E[(X-1)^2]=10,\quad E[(X-2)^2]=6,$$

求 $E(X)$ 与 $D(X)$.

解 $E[(X-1)^2]=E(X^2)-2E(X)+1=10$

$\Longrightarrow E(X^2)-2E(X)=9$,

$E[(X-2)^2]=E(X^2)-4E(X)+4=6$

$\Longrightarrow E(X^2)-4E(X)=2$.

解联立方程组得 $E(X)=7/2$, $E(X^2)=16$. 于是

$$D(X)=E(X^2)-[E(X)]^2=16-\frac{49}{4}=\frac{15}{4}.$$

例 45 已知随机变量 (X,Y) 的联合分布律为

Y \ X	0	1	2
0	0.10	0.25	0.15
1	0.15	0.20	0.15

求: X 和 $X+Y$ 的分布律, $E(Z)=E\left[\sin\frac{\pi(X+Y)}{2}\right]$.

解 X 可取值 0,1,2, $X+Y$ 可取值 0,1,2,3, 故

X	0	1	2
p_k	0.25	0.45	0.30

$X+Y$	0	1	2	3
p_k	0.10	0.40	0.35	0.15

由 $X+Y$ 分布律得

$$\sin 0=\sin\pi=0,\quad \sin\frac{3}{2}\pi=-1,$$

$$E\left[\sin\frac{\pi(X+Y)}{2}\right]=0\times0.10+1\times0.40+0\times0.35+(-1)\times0.15$$

$$=0.25.$$

例 46 设随机变量(X,Y)的概率密度为

$$f(x,y)=\begin{cases}\frac{2}{7}(x+2y), & 0<x<1,1<y<2,\\ 0, & 其它,\end{cases}$$

求

$$E(Z)=E(X/Y^3+X^2Y).$$

解 利用随机变量函数的数学期望公式，有

$$E(Z)=\int_1^2 \mathrm{d}y\int_0^1\left(\frac{x}{y^3}+x^2y\right)\times\frac{2}{7}(x+2y)\mathrm{d}x$$

$$=\int_1^2 \mathrm{d}y\int_0^1 \frac{2}{7}\left(\frac{x^2}{y^3}+\frac{2x}{y^2}+x^3y+2x^2y^2\right)\mathrm{d}x=\frac{46}{63}.$$

例47 射手对目标连续射击，直到命中m 次为止. 设每次射击命中率为p，求所用子弹数X 的数学期望与方差.

解 X 的可取值为$m,m+1,\cdots$. 由于最后一次必然命中，故

$$P\{X=k\}=C_{k-1}^{m-1}p^{m-1}q^{k-m}p,\ k=m,m+1,\cdots$$

(称X 服从帕斯卡分布).

以X_i 记第$i-1$ 次命中至第i 次命中所用子弹数，则

$$X=\sum_{i=1}^m X_i,\ E(X)=\sum_{i=1}^m E(X_i),\ D(X)=\sum_{i=1}^m D(X_i).$$

因为X_i 服从几何分布，$P\{X_i=k\}=q^{k-1}p$，$k=1,2,\cdots$，所以

$$E(X_i)=1/p,\quad D(X_i)=q/p^2$$

(利用级数求和技巧，类似第一节例2 和例27).

于是

$$E(X)=m/p,\quad D(X)=mq/p^2.$$

第二节 其它数字特征

主要内容

1. 协方差与相关系数

(1) 对于二维随机变量(X,Y)，量

$$\{E[X-E(X)][Y-E(Y)]\}$$

称为随机变量 X 与 Y 的协方差，记为 $\operatorname{cov}(X,Y)$，即

$$\operatorname{cov}(X,Y)=E\{[X-E(X)][Y-E(Y)]\}.$$

(2) $\rho_{XY}=\dfrac{\operatorname{cov}(X,Y)}{\sqrt{D(X)}\sqrt{D(Y)}}$称为随机变量 X 与 Y 的相关系数.

(3) 对于任意两个随机变量 X 和 Y，有下列等式成立：

$$D(X+Y)=D(X)+D(Y)+2\operatorname{cov}(X,Y),$$

$$\operatorname{cov}(X,Y)=E(XY)-E(X)E(Y).$$

2. 协方差的性质

(1) $\operatorname{cov}(X,Y)=\operatorname{cov}(Y,X)$；

(2) 对任意常数 a,b，有 $\operatorname{cov}(aX,bY)=ab\cdot\operatorname{cov}(X,Y)$；

(3) $\operatorname{cov}(X_1+X_2,Y)=\operatorname{cov}(X_1,Y)+\operatorname{cov}(X_2,Y)$.

3. 相关系数的两条重要性质

(1) $|\rho_{XY}|\leqslant 1$ ($\rho_{XY}>0$ 称为正相关，$\rho_{XY}<0$ 称为负相关).

(2) $|\rho_{XY}|=1 \Longleftrightarrow$ 存在常数 a,b，使 $P\{Y=a+bX\}=1$ ($|\rho_{XY}|=1$ 称线性相关).

4. 矩(原点矩与中心矩)

设 X 和 Y 是两个随机变量，

(1) 若 $E(X^k)$，$k=1,2,\cdots$ 存在，则称其为 X 的 k 阶原点矩；

(2) 若 $E\{[(X)-E(X)]^k\}$，$k=1,2,\cdots$ 存在，则称其为 X 的 k 阶中心矩；

(3) 若 $E(X^kY^l)$，$k,l=1,2,\cdots$ 存在，则称其为 X 和 Y 的 $k+l$ 阶混合矩；

(4) 若 $E\{[X-E(X)]^k[Y-E(Y)]^l\}$，$k,l=1,2,\cdots$ 存在，则称其为 X 和 Y 的 $k+l$ 阶混合中心矩.

5. 协方差矩阵

n 维随机变量 $(X_1,X_2,\cdots,X_n)$ 的二阶混合中心矩

$$C_{ij}=\operatorname{cov}(X_i,X_j)=E\{[X-E(X)][Y-E(Y)]\}$$

若都存在，$i,j=1,2,\cdots,n$，则称矩阵

$$C=\begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \vdots & \vdots & & \vdots \\ c_{n1} & c_{n2} & \cdots & c_{nn} \end{pmatrix}$$

为n 维随机变量$(X_1,X_2,\cdots,X_n)$的协方差矩阵. 由于$c_{ij}=c_{ji}$ $(i\neq j;i,j=1,2,\cdots,n)$,所以$C$ 是一个对称矩阵.

疑难解析

1. 协方差和相关系数反映了随机变量X与Y之间的什么样的关系?

答 当给定两个随机变量X和Y时,我们必然会考虑它们之间是否存在某种关系. 如当X,Y相互独立时,有

$$D(X+Y)=D(X)+D(Y);$$

一般情况下,

$$D(X+Y)=D(X)+D(Y)+2E\{[X-E(X)][Y-E(Y)]\}.$$

因而量$E\{[X-E(X)][Y-E(Y)]\}\neq 0$ 反映了X,Y不相互独立而存在某种相依关系的事实,将其定义为协方差.

当$D(X),D(Y)$不变时,

$$\rho_{XY}=\operatorname{cov}(X,Y)/[\sqrt{D(X)}\sqrt{D(Y)}]$$

反映了X和Y联系的密切程度. 当$\rho_{XY}=\pm 1$时,X和Y之间存在线性关系$aX+bY+c=0$ 的概率为1;而$\rho_{XY}=0$,只反映X,Y不存在线性关系,不排除其它的联系.

2. 两个随机变量相互独立与不相关有什么区别?怎样解释它们之间的联系?

答 两个随机变量X与Y相互独立与不相关所反映的不是同一种关系. X与Y的独立性反映X与Y之间不存在任何关系,而X与Y不相关只是就线性关系而言的. 但当(X,Y)服从二维正态分布时,X和Y相互独立与X和Y不相关是等价的. 详细地讨论,可

以得出：

(1) 若X,Y相互独立，则X,Y不相关. 因为若X,Y相互独立，则

$$E(XY)=E(X)E(Y),$$

从而 $$\operatorname{cov}(X,Y)=E(XY)-E(X)E(Y)=0,$$

得$\rho_{XY}=0$，所以X,Y不相关.

(2) 若X,Y不相关，则X,Y不一定相互独立. 如例7，由(X,Y)的概率密度

$$f(x,y)=\begin{cases}1/(\pi R^2), & x^2+y^2\leqslant R^2,\\ 0, & \text{其它}\end{cases}$$

可算得 $$E(X)=E(Y)=0,\quad E(X,Y)=0,$$

所以$\rho_{XY}=0$，X与Y不相关. 但

$$f_X(x)=2\sqrt{R^2-x^2}/(\pi R^2),\quad |x|\leqslant R,$$

$$f_Y(y)=2\sqrt{R^2-y^2}/(\pi R^2),\quad |y|\leqslant R,$$

而$f_X(x)f_Y(y)\neq f(x,y)$，所以X与Y不相互独立（例8、例9也说明这一事实）.

(3) 若X,Y相关，则X与Y不相互独立. 因为若X,Y相关，则$\rho_{XY}\neq 0$，即$\operatorname{cov}(X,Y)\neq 0$，因而$E(XY)-E(X)E(Y)\neq 0$，所以$X$与$Y$不相互独立.

(4) 若X,Y不相互独立，X与Y不一定不相关.

方法、技巧与典型例题分析

计算随机变量的其它数字特征的方法与计算随机变量的数学期望和方差的方法一样. 一是依定义计算，但依定义计算时往往要先求随机变量的分布，这样使问题变得复杂；二是由随机变量与数字特征之间的关系来计算，这样比较灵活，也比较简捷，但对概念的熟悉与技巧的熟练程度要求较高.

一、其它数字特征的计算

例 1 设 $D(X)=4, D(Y)=9, \rho_{XY}=0.6$，则 $D(3X-2Y)=$ ______.

解
$$D(3X-2Y)=D(3X)+D(2Y)-2\rho_{XY}\times 2\times 3\sqrt{D(X)}\sqrt{D(Y)}$$
$$=9\times 4+4\times 9-12\times 0.6\times 2\times 3$$
$$=28.8.$$

例 2 已知随机变量 X 的方差 $D(X)$ 有限，设 $Y=aX+b$ ($a\neq 0, a, b$ 为常数)，则 $\rho_{XY}=$（　　）.

(A) 1;　　(B) -1;　　(C) $a/|a|$;　　(D) $|\rho_{XY}|<1$.

解 选(C). 因为

$$E(XY)=E[X(aX+b)]=aE(X^2)+bE(X),$$
$$\mathrm{cov}(X,Y)=E(XY)-E(X)E(Y)$$
$$=aE(X^2)+bE(X)-E(X)[aE(X)+b]$$
$$=aE(X^2)-a[E(X)]^2=aD(X),$$

而
$$D(Y)=a^2D(X),\quad \sqrt{D(Y)}=|a|\sqrt{D(X)},$$

所以
$$\rho_{XY}=\frac{aD(X)}{|a|\sqrt{D(X)}\sqrt{D(X)}}=\frac{a}{|a|}.$$

例 3 若随机变量 X 与 Y 满足 $D(X+Y)=D(X-Y)$，则必有（　　）.

(A) X 与 Y 相互独立;　　(B) X 与 Y 不相关;

(C) $D(Y)=0$;　　(D) $D(X)D(Y)=0$.

解 选(B). 因为

$$D(X+Y)=D(X)+D(Y)+2\mathrm{cov}(X,Y),$$
$$D(X-Y)=D(X)+D(Y)-2\mathrm{cov}(X,Y),$$

所以由 $D(X+Y)=D(X-Y)$ 得 $\mathrm{cov}(X,Y)=0\Longrightarrow\rho_{XY}=0$，从而知 X 与 Y 不相关.

例 4 设对随机变量 X, Y 有 $E(X)=2, E(X^2)=20, E(Y)=3, E(Y^2)=34, \rho_{XY}=0.5$，求:

(1) $E(3X+2Y), D(3X+2Y)$；

(2) $E(X-Y), D(X-Y)$.

解 (1) $E(3X+2Y)=3E(X)+2E(Y)=6+6=12$，

$$\begin{aligned}D(3X+2Y)&=9D(X)+4D(Y)+2\mathrm{cov}(X,Y)\\&=9\times(20-4)+4\times(34-9)\\&\quad+12\times0.5\times4\times5=364.\end{aligned}$$

(2) $E(X-Y)=E(X)-E(Y)=-1$，

$$\begin{aligned}D(X-Y)&=D(X)+D(Y)-2\mathrm{cov}(X,Y)\\&=(20-4)+(34-9)-2\times0.5\times4\times5=21.\end{aligned}$$

例 5 设随机变量(X,Y)的协方差矩阵为

$$\boldsymbol{C}=\begin{pmatrix}4 & -3\\ -3 & 9\end{pmatrix},$$

求X和Y的相关系数ρ_{XY}.

解 由协方差矩阵知

$$D(X)=4,\quad D(Y)=9,\quad \mathrm{cov}(X,Y)=-3,$$

故
$$\rho_{XY}=-3/(2\times3)=-1/2.$$

例 6 设随机变量(X,Y)在区域$D: x^2+y^2\leqslant R^2$上服从均匀分布. (1)问X,Y是否相互独立； (2)求X,Y的相关系数ρ_{XY}.

解 (1) 因为

$$f(x,y)=\begin{cases}1/(\pi R^2), & x^2+y^2<R^2,\\ 0, & \text{其它},\end{cases}$$

所以

$$f_X(x)=\begin{cases}\dfrac{1}{\pi R^2}\displaystyle\int_{-\sqrt{R^2-x^2}}^{\sqrt{R^2-x^2}}\mathrm{d}y=\dfrac{2}{\pi R^2}\sqrt{R^2-x^2}, & |x|\leqslant R,\\ 0, & \text{其它}.\end{cases}$$

类似地，
$$f_Y(y)=\begin{cases}\dfrac{1}{\pi R^2}\sqrt{R^2-y^2}, & |y|\leqslant R,\\ 0, & \text{其它}.\end{cases}$$

显然，$f_X(x)f_Y(y)\neq f(x,y)$，所以X与Y不相互独立.

(2) $E(X)=\frac{2}{\pi R^2}\int_{-R}^{R}x\sqrt{R^2-x^2}\mathrm{d}x \xlongequal{\text{奇偶性}} 0,\quad E(Y)=0,$

同样 $$\mathrm{cov}(X,Y)=E(XY)=\frac{1}{\pi R^2}\iint\limits_{x^2+y^2\leqslant R^2}xy\mathrm{d}x\mathrm{d}y=0,$$

所以 $\rho_{XY}=0$，即 X,Y 不相关. 由此可见，不相关不一定相互独立.

例 7 设随机变量 X 服从拉普拉斯分布，其概率密度为

$$f(x)=\frac{1}{2}\mathrm{e}^{-|x|},\quad -\infty<x<+\infty,$$

求：(1) $E(|X|),D(|X|)$；

(2) $\mathrm{cov}(X,|X|)$，问：X 与 $|X|$ 是否相互独立？是否不相关？

解 (1) $E(|X|)=\frac{1}{2}\int_{-\infty}^{+\infty}|x|\mathrm{e}^{-|x|}\mathrm{d}x \xlongequal{\text{偶}} \int_0^{+\infty}x\mathrm{e}^{-x}\mathrm{d}x$

$=\Gamma(2)=1,$

$$E(|X|^2)=\frac{1}{2}\int_{-\infty}^{+\infty}|x|^2\mathrm{e}^{-|x|}\mathrm{d}x \xlongequal{\text{偶}} \int_0^{x}x^2\mathrm{e}^{-x}\mathrm{d}x=\Gamma(3)=2,$$

故 $$D(|X|)=E(|X|^2)-[E(|X|)]^2=2-1=1.$$

(2) $E(X|X|)=\int_{-\infty}^{+\infty}\frac{x}{2}|x|\mathrm{e}^{-|x|}\mathrm{d}x$

$=\frac{-1}{2}\int_{-\infty}^{0}x^2\mathrm{e}^{x}\mathrm{d}x+\frac{1}{2}\int_0^{+\infty}x^2\mathrm{e}^{-x}\mathrm{d}x$

$=-\frac{1}{2}\Gamma(3)+\frac{1}{2}\Gamma(3)=0,$

$E(X)=\frac{1}{2}\int_{-\infty}^{+\infty}x\mathrm{e}^{-|x|}\mathrm{d}x \xlongequal{\text{奇}} 0,$

故 $$\mathrm{cov}(X,Y)=E(XY)-E(X)E(Y)=0-1\times 0=0.$$

所以，X 与 $|X|$ 不相关.

对于 $0<a<\infty$，$\{|X|<a\}\subset\{X<a\}$. 由于 $P\{|X|<a\}>0$，$P\{X<a\}<1$，所以

$$P\{X<a\}=P\{|X|<a\}\neq P\{|X|<a\}P\{X<a\}.$$

从而知 X 与 Y 不相互独立.

例 8 已知随机变量 (X,Y) 的分布律为

$Y \backslash X$	-1	0	1
-1	1/8	1/8	1/8
0	1/8	0	1/8
1	1/8	1/8	1/8

试验证 X 与 Y 不相关，但 X 与 Y 不相互独立.

证 $E(X)=-1\times 3/8+0\times 2/8+1\times 3/8=0$,

$$E(X^2)=1\times 3/8+0\times 2/8+1\times 3/8=3/4,$$

$$D(X)=E(X^2)-[E(X)]^2=3/4,$$

$$E(Y)=0,\quad D(Y)=3/4.$$

而 $$E(X,Y)=1/8-1/8-1/8+1/8=0,$$

所以 $$\operatorname{cov}(X,Y)=E(XY)-E(X)E(Y)=0,$$

故 $\rho_{XY}=0$，即 X 与 Y 不相关.

任取表中 p_{ij}，如 $p_{11}=1/8$，而 $p_{1\cdot}=3/8$，$p_{\cdot 1}=3/8$，显然 $p_{11}\neq p_{1\cdot}p_{\cdot 1}$，所以 X 与 Y 不相互独立.

例 9 已知随机变量 $X\sim N(0,1)$，$Y=X^3$，求 X 与 Y 的相关系数.

解 由上节例 19 知，若 $X\sim N(0,1)$，则

$$E(X^k)=\begin{cases}(k-1)!!, & k \text{ 为偶数},\\ 0, & k \text{ 为奇数}.\end{cases}$$

所以 $$E(X)=0,\quad D(X)=1,$$

$$E(Y)=0,\quad E(Y^2)=E(X^6)=15,\quad D(Y)=15.$$

又 $$\operatorname{cov}(X,Y)=E(XY)-E(X)E(Y)=E(X^4)-0=3,$$

所以 $$\rho_{XY}=\frac{\operatorname{cov}(X,Y)}{\sqrt{D(X)}\sqrt{D(Y)}}=\frac{3}{1\times\sqrt{15}}=\frac{1}{5}\sqrt{15}.$$

例 10 已知随机变量 X,Y,Z，证明：

$$D(X+Y+Z)=D(X)+D(Y)+D(Z)+2\operatorname{cov}(X,Y)+2\operatorname{cov}(X,Z)+2\operatorname{cov}(Y,Z).$$

证 由方差定义展开上式可得

$$
\begin{aligned}
&D(X+Y+Z)\\
&=E\{[(X+Y+Z)-E(X+Y+Z)]^2\}\\
&=E\{[X-E(X)]+[Y-E(Y)]+[Z-E(Z)]^2\}\\
&=E\{[X-E(X)]^2\}+E\{[Y-E(Y)]^2\}+E\{[Z-E(Z)]^2\}\\
&\quad+2E\{[X-E(X)][Y-E(Y)]\}\\
&\quad+2E\{[X-E(X)][Z-E(Z)]\}\\
&\quad+2E\{[Y-E(Y)][Z-E(Z)]\}\\
&=D(X)+D(Y)+D(Z)+2\operatorname{cov}(X,Y)\\
&\quad+2\operatorname{cov}(X,Z)+2\operatorname{cov}(Y,Z).
\end{aligned}
$$

例 11 已知随机变量(X,Y,Z)的协方差矩阵

$$
\boldsymbol{C}=\begin{pmatrix} 9 & 1 & -2 \\ 1 & 20 & 3 \\ -2 & 3 & 12 \end{pmatrix},
$$

令 $X_1=2X+3Y+Z, Y_1=X-2Y+5Z, Z_1=Y-Z$，求$(X_1,Y_1,Z_1)$的协方差矩阵.

解 由协方差矩阵知，$D(X)=9, D(Y)=20, D(Z)=12$，$\operatorname{cov}(X,Y)=1, \operatorname{cov}(X,Z)=-2, \operatorname{cov}(Y,Z)=3$. 利用上题结论和协方差运算性质，得

$$
\begin{aligned}
D(X_1)&=4D(X)+9D(Y)+D(Z)+12\operatorname{cov}(X,Y)\\
&\quad+4\operatorname{cov}(X,Z)+6\operatorname{cov}(Y,Z)\\
&=36+180+12+12-8+18=250,\\
D(Y_1)&=D(X)+4D(Y)+25D(Z)\\
&\quad+(-4)\operatorname{cov}(X,Y)+10\operatorname{cov}(X,Z)\\
&\quad+(-20)\operatorname{cov}(Y,Z)=305.\\
D(Z_1)&=D(Y)+D(Z)-2\operatorname{cov}(Y,Z)=26,\\
\operatorname{cov}(X_1,Y_1)&=\operatorname{cov}(2X+3Y+Z, X-2Y+5Z)=-26.
\end{aligned}
$$

同理 $\qquad \operatorname{cov}(X_1,Z_1)=48, \quad \operatorname{cov}(Y_1,Z_1)=-76.$

(X_1,Y_1,Z_1)的协方差矩阵为$\begin{pmatrix} 250 & -26 & 48 \\ -26 & 305 & -76 \\ 48 & -76 & 26 \end{pmatrix}$.

例12 设随机变量(X,Y)的概率密度为

$$f(x,y)=\begin{cases}1, & |y|\leqslant x, 0\leqslant x\leqslant 1,\\ 0, & 其它,\end{cases}$$

求:(1) $f_X(x), f_Y(y)$; (2) $E(X), E(Y), D(X), D(Y)$;

(3) $\operatorname{cov}(X,Y)$.

解 (1)为了易于分析积分区域,画出$f(x,y)$不为零的区域图形(见图4.5).

图4.5

当$0\leqslant x\leqslant 1$时,

$$f_X(x)=\int_{-x}^{x}1\mathrm{d}y=2x,$$

所以 $$f_X(x)=\begin{cases}2x, & 0\leqslant x\leqslant 1,\\ 0, & 其它;\end{cases}$$

当$0<y<1$时, $f_Y(y)=\int_{y}^{1}1\mathrm{d}x=1-y$;

当$-1<y<0$时, $f_Y(y)=\int_{-y}^{1}1\mathrm{d}x=1+y$.

所以 $$f_Y(y)=\begin{cases}1-|y|, & -1\leqslant y\leqslant 1,\\ 0, & 其它.\end{cases}$$

又 $$E(X)=\int_0^1 x\times 2x\mathrm{d}x=\frac{2}{3},$$

$$E(X^2)=\int_0^1 x^2\times 2x\mathrm{d}x=\frac{1}{2},$$

$$E(Y)=\int_{-1}^{1}(1-|y|)\mathrm{d}y\xlongequal{奇}0,$$

$$E(Y^2)=2\int_0^1 y^2(1-y)\mathrm{d}y=\frac{1}{6},$$

所以 $$D(X)=\frac{1}{2}-\left(\frac{2}{3}\right)^2=\frac{1}{18},\quad D(Y)=\frac{1}{6},$$

$$E(XY)=\int_0^1\mathrm{d}x\int_{-x}^{x}xy\mathrm{d}y\xlongequal{奇}0,$$

故 $$\mathrm{cov}(X,Y)=E(X)E(Y)=0.$$

例 13 已知 $X\sim N(1,3^2)$，$Y\sim N(0,4^2)$，$\rho_{XY}=-1/2$，对 $Z=X/3+Y/2$，求：

(1) $E(Z)$和$D(Z)$； (2) ρ_{XZ}； (3) X,Z 的独立性.

解 (1) 由已给分布知

$$E(X)=1,\quad D(X)=9,\quad E(Y)=0,\quad D(Y)=16,$$

所以 $$E(Z)=\frac{1}{3}E(X)+\frac{1}{2}E(Y)=\frac{1}{3}+0=\frac{1}{3},$$

$$D(Z)=\frac{1}{9}D(X)+\frac{1}{4}D(Y)+2\times\frac{1}{3}\times\frac{1}{2}\mathrm{cov}(X,Y)$$
$$=1+4+\frac{1}{3}\left(-\frac{1}{2}\right)\times 4\times 3=3.$$

(2) $$\mathrm{cov}(X,Z)=\mathrm{cov}\left(X,\frac{X}{3}+\frac{Y}{2}\right)$$
$$=\frac{1}{3}\mathrm{cov}(X,X)+\frac{1}{2}\mathrm{cov}(X,Y)$$
$$=\frac{1}{3}D(X)+\frac{1}{2}\left(-\frac{1}{2}\right)\times 3\times 4=3-3=0,$$

所以 $$\rho_{XZ}=0.$$

(3) 因为$Z=\frac{X}{3}+\frac{Y}{2}$，所以$Z$ 也是正态分布. 对于正态分布，不相关与相互独立等价，故由 $\rho_{XZ}=0$ 知，X,Z 相互独立.

例 14 设二维随机变量$(X,Y)\sim N(0,0,\sigma_1^2,\sigma_2^2,\rho)$，其中 $\sigma_1^2\neq\sigma_2^2$. 又设 $X_1=X\cos\alpha+Y\sin\alpha$，$X_2=-X\sin\alpha+Y\cos\alpha$，问：何时 X_1 与 X_2 不相关，且 X_1 与 X_2 相互独立？

解 因为(X_1,X_2)是(X,Y)的线性变换，所以(X_1,X_2)仍然是二维正态随机变量，若 X_1 与 X_2 不相关，X_1 与 X_2 必然相互独立.

$$E(X_1)=E(X_2)=0,$$

$$\mathrm{cov}(X_1,X_2)=E[(X\cos\alpha+Y\sin\alpha)(-X\sin\alpha+Y\cos\alpha)]-0$$
$$=E[-X^2\sin\alpha\cos\alpha+Y^2\sin\alpha\cos\alpha+XY(\cos^2\alpha-\sin^2\alpha)]$$
$$=(\sigma_1^2-\sigma_2^2)\sin\alpha\cos\alpha+\rho\sigma_1\sigma_2(\cos^2\alpha-\sin^2\alpha).$$

若 X_1 与 X_2 不相关，则 $\mathrm{cov}(X_1,X_2)=0$. 从而有

$$\tan 2\alpha=\frac{2\sin\alpha\cos\alpha}{\cos^2\alpha-\sin^2\alpha}=\frac{2\rho\sigma_1\sigma_2}{\sigma_1^2-\sigma_2^2}.$$

此时，X_1 与 X_2 不相关，且 X_1 与 X_2 相互独立.

例 15 设随机变量 (X,Y) 的概率密度为

$$f(x,y)=\begin{cases}2-x-y, & 0<x,y<1,\\ 0, & \text{其它},\end{cases}$$

求 (X,Y) 的协方差矩阵和相关矩阵.

解 $E(X)=\int_0^1 \mathrm{d}x\int_0^1 x(2-x-y)\mathrm{d}y=\frac{5}{12}$，$E(Y)=\frac{5}{12}$，

$$E(X^2)=\int_0^1 \mathrm{d}x\int_0^1 x^2(2-x-y)\mathrm{d}y=\frac{1}{4},\quad E(Y^2)=\frac{1}{4},$$

$$D(X)=E(X^2)-[E(X)]^2=\frac{1}{4}-\left(\frac{5}{12}\right)^2=\frac{11}{144},\quad D(Y)=\frac{11}{144},$$

$$E(XY)=\int_0^1 \mathrm{d}x\int_0^1 xy(2-x-y)\mathrm{d}y=\frac{1}{6},$$

$$\mathrm{cov}(X,Y)=E(X,Y)-E(X)E(Y)=\frac{1}{6}-\left(\frac{5}{12}\right)^2=-\frac{1}{144}.$$

又

$$\rho_{XX}=\rho_{YY}=1,$$

$$\rho_{XY}=\frac{\mathrm{cov}(X,Y)}{\sqrt{D(X)}\sqrt{D(Y)}}=-\frac{1}{11}.$$

所以协方差矩阵 $\boldsymbol{C}$ 与相关系数矩阵 $\boldsymbol{P}$ 分别为

$$\boldsymbol{C}=\begin{pmatrix}\frac{11}{144} & -\frac{1}{144}\\ -\frac{1}{144} & \frac{11}{144}\end{pmatrix},\quad \boldsymbol{P}=\begin{pmatrix}1 & -\frac{1}{11}\\ -\frac{1}{11} & 1\end{pmatrix}.$$

例 16 设 X 和 Y 都是标准化随机变量，$\rho_{XY}=1/2$，令 $Z_1=aX$，$Z_2=bX+cY$，试确定 a,b,c 的值，使 $D(Z_1)=D(Z_2)=1$，且 Z_1 与 Z_2 不相关.

解 因为 X,Y 都是标准化随机变量，即

$$E(X)=0,\quad E(Y)=0,\quad D(X)=1,\quad D(Y)=1.$$

又 $\rho_{XY}=1/2$，于是

$$D(Z_1)=D(aX)=a^2D(X)=a^2,$$

$$D(Z_2)=D(bX+cY)=b^2D(X)+c^2D(Y)+2\times 1/2\times bc\times 1\times 1$$
$$=b^2+c^2+bc,$$

$$\operatorname{cov}(X,Y)=\rho_{XY}\sqrt{D(X)}\sqrt{D(Y)}=1/2,$$

$$\operatorname{cov}(Z_1,Z_2)=\operatorname{cov}(aX,bX+cY)$$
$$=ab\operatorname{cov}(X,X)+ac\operatorname{cov}(X,Y)=ab+bc/2.$$

建立方程组$\begin{cases}a^2=1,\\ b^2+c^2+bc=1,\\ ab+bc/2=0,\end{cases}$解得$\begin{cases}a=\pm 1,\\ b=\pm 1/\sqrt{3},\\ c=\mp 2/\sqrt{3}.\end{cases}$

例 17 设$X_1,X_2,\cdots,X_{n+m}$ $(n>m)$是相互独立且同分布、而且方差存在的随机变量，又令$Y=X_1+X_2+\cdots+X_n$，$Z=X_{m+1}+X_{m+2}+\cdots+X_{m+n}$，求$\rho_{YZ}$.

解 为简单计，不妨设

$$E(X_i)=0,\quad E(X_i^2)=1,\quad i=1,2,\cdots,m+n,$$

则$\operatorname{cov}(Y,Z)=E[(X_1+X_2+\cdots+X_n)(X_{m+1}+X_{m+2}+\cdots+X_{m+n})]$

$$=E[X_{m+1}^2+X_{m+2}^2+\cdots+X_n^2]=n-m,$$

$$D(Y)=E(X_1^2+X_2^2+\cdots+X_n^2)=n,$$

$$D(Z)=E(X_{m+1}^2+X_{m+2}^2+\cdots+X_{m+n}^2=n,$$

故
$$\rho_{XY}=\frac{\operatorname{cov}(Y,Z)}{\sqrt{D(Y)}\sqrt{D(Z)}}=\frac{n-m}{n}=1-\frac{m}{n}.$$

例 18 设$X_1,X_2,\cdots,X_{2n}$的数学期望为零，方差为1，且任何两个随机变量的相关系数为ρ. 令$Y=X_1+X_2+\cdots+X_n$，$Z=X_{n+1}+X_{n+2}+\cdots+X_{2n}$，求$\rho_{YZ}$.

解

$$\operatorname{cov}(Y,Z)=E[(X_1+X_2+\cdots+X_n)(X_{n+1}+X_{n+2}+\cdots+X_{2n})]$$
$$=n^2\rho.$$

$$D(Y)=E(X_1+X_2+\cdots+X_n)^2$$
$$=E(X_1^2+X_2^2+\cdots+X_n^2)+2\sum_{1\leqslant i<j\leqslant n}E(X_iX_j)$$

$$=n+2C_n^2\rho=n+n(n-1)\rho=n[1+(n-1)\rho],$$

$$D(Z)=n+n(n-1)\rho=n[1+(n-1)\rho].$$

所以 $$\rho_{YZ}=\frac{\operatorname{cov}(Y,Z)}{\sqrt{D(Y)}\sqrt{D(Z)}}=\frac{n\rho}{1+(n-1)\rho}.$$

例 19 设随机变量(X,Y)的概率密度为

$$f(x,y)=\begin{cases}k\cos(x+y), & 0\leqslant x\leqslant \pi/2,-\pi/2\leqslant y\leqslant 0,\\ 0, & \text{其它},\end{cases}$$

求$E(X),E(Y),\operatorname{cov}(X,Y),\rho_{XY}$.

解 由

$$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}f(x,y)\mathrm{d}x\mathrm{d}y=\int_0^{\pi/2}\mathrm{d}x\int_{-\pi/2}^{0}k\cos(x+y)\mathrm{d}y=2k=1$$

得 $$k=1/2,$$

所以 $$E(X)=\frac{1}{2}\int_0^{\pi/2}x\mathrm{d}x\int_{-\pi/2}^{0}\cos(x+y)\mathrm{d}y=-0.785,$$

$$D(X)=\frac{1}{2}\int_0^{\pi}x^2\mathrm{d}x\int_{-\pi/2}^{0}\cos(x+y)\mathrm{d}y-[E(X)]^2=0.188.$$

类似地， $$E(Y)=-0.785,\quad D(Y)=0.188.$$

又 $$E(XY)=\frac{1}{2}\int_0^{\pi/2}x\mathrm{d}x\int_{-\pi/2}^{0}y\cos(x+y)\mathrm{d}y=-0.663,$$

故 $$\operatorname{cov}(X,Y)=E(XY)-E(X)E(Y)$$

$$=-0.663+(0.785)^2=-0.046,$$

$$\rho_{XY}=\frac{\operatorname{cov}(X,Y)}{\sqrt{D(X)}\sqrt{D(Y)}}=-0.244.$$

例20 设随机变量$X\sim U[0,2]$，求X与$|X-1|$的相关系数与协方差矩阵$\boldsymbol{C}$.

解 因为 $$f_X(x)=\begin{cases}1/2, & 0\leqslant x\leqslant 2,\\ 0, & \text{其它},\end{cases}$$

所以 $$E(X)=(2+0)/2=1,\quad D(X)=(2-0)^2/12=1/3,$$

$$E(|X-1|)=\frac{1}{2}\int_0^2|x-1|\mathrm{d}x$$

$$=\frac{1}{2}\left[\int_0^1(1-x)\mathrm{d}x+\int_1^2(x-1)\mathrm{d}x\right]$$

$$=\frac{1}{4}\left[-(x-1)^2\Big|_0^1+(x-1)^2\Big|_1^2\right]=\frac{1}{2},$$

$$D(|X-1|)=E[(X-1)^2]-[E(|X-1|)]^2$$

$$=D(X)-\frac{1}{4}=\frac{1}{12}.$$

而 $$E[X|X-1|]=\frac{1}{2}\int_0^2 x|x-1|\mathrm{d}x$$

$$=-\frac{1}{2}\int_0^1(x^2-x)\mathrm{d}x+\frac{1}{2}\int_1^2(x^2-x)\mathrm{d}x=\frac{1}{2},$$

所以 $$\mathrm{cov}(X,|X-1|)=\frac{1}{2}-\frac{1}{2}=0.$$

于是 $$\rho_{X|X-1|}=\frac{\mathrm{cov}(X,|X-1|)}{\sqrt{D(X)}\sqrt{D(|X-1|)}}=0,$$

协方差矩阵 $$\boldsymbol{C}=\frac{1}{12}\begin{pmatrix}4 & 0\\ 0 & 1\end{pmatrix}.$$

例 21 设随机变量 X 的分布律为

X	-2	-1	1	2
p_k	$1/4$	$1/4$	$1/4$	$1/4$

验证 X^2 与 X 不相关，X^3 与 X 相关.

证 X^2，X^3，X^4 的分布律分别为

X^2	1	4
p_k	$1/2$	$1/2$

X^3	-8	-1	1	8
p_k	$1/4$	$1/4$	$1/4$	$1/4$

X^4	1	16
p_k	$1/2$	$1/2$

于是 $E(X)=0$，$E(X^2)=5/2$，$E(X^3)=0$，$E(X^4)=17/2$，

$$\mathrm{cov}(X,X^2)=E(X^3)-E(X)E(X^2)=0\Longrightarrow\rho_{XX^2}=0,$$

$$\mathrm{cov}(X,X^3)=E(X^4)-E(X)E(X^3)=17/2\Longrightarrow\rho_{XX^3}\neq 0.$$

所以 X 与 X^2 不相关，X 与 X^3 相关.

例 22 设随机变量 (X,Y) 的概率密度为

$$f(x,y)=\begin{cases}\frac{1}{4}\sin x\sin y, & 0\leqslant x,y\leqslant\pi,\\ 0, & \text{其它},\end{cases}$$

求 $E(X,Y)$，$D(X,Y)$ 及 ρ_{XY}.

解 由 $\int_0^\pi \frac{1}{4}\sin x\sin y\mathrm{d}y=\frac{1}{2}\sin x$，知

$$f_X(x)=\begin{cases}\frac{1}{2}\sin x, & 0\leqslant x\leqslant\pi,\\ 0, & \text{其它},\end{cases}\qquad f_Y(y)=\begin{cases}\frac{1}{2}\sin y, & 0\leqslant y\leqslant\pi,\\ 0, & \text{其它}.\end{cases}$$

显然，有 $f_X(x)f_Y(y)=f(x,y)$，所以 X,Y 相互独立，$\rho_{XY}=0$.

$$E(X)=\int_0^\pi \frac{x}{2}\sin x\mathrm{d}x=\frac{\pi}{2},\quad E(Y)=\int_0^\pi \frac{y}{2}\sin y\mathrm{d}y=\frac{\pi}{2},$$

$$E(X^2)=\int_0^\pi \frac{x^2}{2}\sin x\mathrm{d}x=\frac{\pi^2}{2}-2,\quad E(Y^2)=\frac{\pi^2}{2}-2,$$

$$D(X)=E(X^2)-[E(X)]^2=\frac{\pi^2}{4}-2,\quad D(Y)=\frac{\pi^2}{4}-2.$$

所以 $E(X,Y)=\left(\frac{\pi}{2},\frac{\pi}{2}\right)$，$D(X,Y)=\left(\frac{\pi^2}{4}-2,\frac{\pi^2}{4}-2\right)$.

例 23 设随机变量 $X\sim U[0,2\pi]$，概率密度

$$f_X(x)=\begin{cases}1/(2\pi), & 0\leqslant x\leqslant 2\pi,\\ 0, & \text{其它}.\end{cases}$$

令 $Y=\sin X$，$Z=\sin(x+a)$，$a\in[0,2\pi]$ 为常数.

(1) 求 ρ_{YZ}； (2) 讨论 Y 和 Z 的相关性与独立性.

解
$$E(Y)=\int_0^{2\pi}\sin x\times\frac{1}{2\pi}\mathrm{d}x=0,$$

$$E(Z)=\int_0^{2\pi}\sin(x+a)\times\frac{1}{2\pi}\mathrm{d}x=0,$$

$$E(Y^2)=\int_0^{2\pi}\sin^2 x\,\frac{1}{2\pi}\mathrm{d}x=\frac{1}{4\pi}\int_0^{2\pi}(1-\cos 2x)\mathrm{d}x=\frac{1}{2},$$

$$E(Z^2)=\int_0^{2\pi}\sin^2(x+a)\frac{1}{2\pi}\mathrm{d}x=\frac{1}{4\pi}\int_0^{2\pi}[1-\cos 2(x+a)]\mathrm{d}x=\frac{1}{2}.$$

所以 $D(Y)=E(Y^2)-[E(Y)]^2=1/2$，$D(Z)=1/2$.

$$E(YZ)=\int_0^{2\pi}\sin x\sin(x+a)\frac{1}{2\pi}\mathrm{d}x$$
$$=\frac{1}{4\pi}\int_0^{2\pi}[\cos a-\cos(2x+a)]\mathrm{d}x=\frac{1}{2}\cos a,$$

故 $$\mathrm{cov}(Y,Z)=E(YZ)-0=\frac{1}{2}\cos a,$$

$$\rho_{YZ}=\frac{\mathrm{cov}(Y,Z)}{\sqrt{D(Y)}\sqrt{D(Z)}}=\cos a.$$

讨论：当 $a=0,\pi,2\pi$ 时，$|\rho_{YZ}|=1$，Y 与 Z 是否存在线性关系？

当 $a=0,2\pi$ 时，

$$\rho_{YZ}=1,\quad \sin(X+a)=\sin X\ (Z=Y);$$

当 $a=\pi$ 时，

$$\rho_{YZ}=-1,\quad \sin(X+a)=-\sin X\ (Z=-Y).$$

当 $a=\frac{\pi}{2},\frac{3}{2}\pi$ 时，$\rho_{YZ}=0$，Y 与 Z 不相关. 但 Y 与 Z 不相互独立，因为 Y 与 Z 有以下关系

$$Y^2+Z^2=\sin^2(X+a)+\sin^2X=\cos^2X+\sin^2X=1.$$

例 24 一个工人照看 n 台机床. 设 n 台机床排成一行，相邻两台机床的间距为 a，求该工人从已照看机床到待照看的下一台的机床间的行走距离的数学期望.

解 将机床顺次编号为 $1,2,\cdots,n$，设工人已照看的机床为第 k 台，每台机床需照看的概率相同，均为 $\frac{1}{n}$. 工人需走的距离为

$$d_{ki}=\begin{cases}(k-i)a, & k\geqslant i,\\ (i-k)a, & k<i,\end{cases}\quad 0\leqslant i\leqslant n.$$

所以 $$E(d_{ki}|k)=\frac{1}{n}\left[\sum_{i=1}^{k}(k-i)a+\sum_{i=k+1}^{n}(i-k)a\right]$$

$$=\frac{a}{n}\left(\sum_{u=0}^{k-1}u+\sum_{v=1}^{n-k}v\right)$$

$$=\frac{a}{2n}[2k^2-2(n+1)k+n(n+1)].$$

故工人在各车床间行走距离的期望值为

$$E(d)=E[E(d_{ki}|k)]=\sum_{k=1}^{n}P(k)E[d_{ki}|k]$$

$$=\sum_{k=1}^{n}\frac{a}{2n^2}[2k^2-2(n+1)k+n(n-1)]$$

$$=\frac{a}{3n}(n^2-1).$$

二、关于数字特征的证明题

数字特征的证明题,反映了数字特征的特性、数字特征之间的相互关系以及一些可以作为公式应用或推广的结论. 在做证明题时,首先要求读者对数字特征的概念和性质十分熟悉,其次要熟谙多种证题的技巧. 做证明题常用的技巧有:(1)寻找不同概念之间的联系,由联系入手,步步推进;(2)用"拆拼"项的手段,将原来的项组合成新项,完成证明;(3)由于数字特征由级数和积分计算,因此要善于用级数求和与积分求积的技巧来证明;(4)利用对称性、独立性、奇偶性、互逆性证明.

在做证明题时最重要还是要有开阔的思路,只有进行广泛的联想,才能灵活地运用技巧,所以读者必须学会多看、多想、多实践.

例 25 设随机变量 X 的概率密度为

$$f(x)=\frac{1}{m!}x^m e^{-x} \quad (x\geqslant 0, m \text{ 为正整数}),$$

证明:
$$E(X)=D(X)=m+1.$$

证 这是一道利用计算结果来证的证明题,关键是运用了 Γ 函数的性质

$$\Gamma(m+1)=\int_0^{+\infty}x^m e^{-x}\mathrm{d}x=m!.$$

因为 $E(X)=\frac{1}{m!}\int_0^{+\infty}x\cdot x^m e^{-x}\mathrm{d}x=\frac{1}{m!}(m+1)!=m+1,$

$$E(X^2)=\frac{1}{m!}\int_0^{+\infty}x^2x^m e^{-x}\mathrm{d}x=\frac{1}{m!}(m+2)!=(m+2)(m+1),$$

所以
$$\begin{aligned}D(X)&=E(X^2)-[E(X)]^2\\&=(m+2)(m+1)-(m+1)^2\\&=m+1=E(X).\end{aligned}$$

例 26 设随机变量 X 的 $E(X)$，$D(X)$ 均存在，且 $D(X)$ 不等于零，标准化随机变量为 $Y=\dfrac{X-E(X)}{\sqrt{D(X)}}$，证明随机变量 $Y\sim N(0,1)$.

证 利用数学期望与方差的性质来证.

$$E(Y)=E\left(\frac{X-E(X)}{\sqrt{D(X)}}\right)=\frac{1}{\sqrt{D(X)}}E[X-E(X)]$$

$$=\frac{1}{\sqrt{D(X)}}[E(X)-E(X)]=0,$$

$$D(Y)=D\left(\frac{X-E(X)}{\sqrt{D(X)}}\right)=\frac{1}{D(X)}D[X-E(X)]\ (E(X)\text{是常数})$$

$$=\frac{1}{D(X)}D(X)=1,$$

所以，$Y\sim N(0,1)$.

例 27 设 X 为取值于 (a,b) 的连续型随机变量. 证明：

(1) $a\leqslant E(X)\leqslant b$； (2) $D(X)\leqslant(b-a)^2/4$.

证 (1) 因为在 (a,b) 内，$f(x)\neq 0$，所以

$$E(X)=\int_a^b xf(x)\mathrm{d}x\geqslant a\int_a^b f(x)\mathrm{d}x=a\times 1=a,$$

$$E(X)=\int_a^b xf(x)\mathrm{d}x\leqslant b\int_a^b f(x)\mathrm{d}x=b\times 1=b,$$

于是 $$a\leqslant E(X)\leqslant b.$$

(2) 由 $a\leqslant X\leqslant b$，得

$$-\frac{b-a}{2}\leqslant X-\frac{a+b}{2}\leqslant b-\frac{a+b}{2}=\frac{b-a}{2},$$

即 $$\left(X-\frac{a+b}{2}\right)^2\leqslant\left(\frac{b-a}{2}\right)^2\Longrightarrow E\left[\left(X-\frac{a+b}{2}\right)^2\right]\leqslant\left(\frac{b-a}{2}\right)^2.$$

于是 $$D(X)=E\{[X-E(X)]^2\}$$

$$=E\left(\left\{\left[X-\frac{a+b}{2}\right]+\left[\frac{a+b}{2}-E(X)\right]\right\}^2\right)$$

$$=E\left[\left(X-\frac{a+b}{2}\right)^2\right]-\left[\frac{a+b}{2}-E(X)\right]^2$$

$$\leqslant E\left[\left(X-\frac{a+b}{2}\right)^2\right]\leqslant\frac{(b-a)^2}{4}.$$

例 28 设 X 为取非负整数值的随机变量. 证明:

$$E(X)=\sum_{n=1}^{\infty}P\{X\geqslant n\}.$$

证 利用级数交换和号的技巧,则

$$E(X)=\sum_{k=1}^{\infty}kP\{X=k\}=\sum_{k=1}^{\infty}\sum_{n=1}^{k}P\{X=k\}$$

$$=\sum_{n=1}^{\infty}\sum_{k=n}^{\infty}P\{X=k\}=\sum_{n=1}^{\infty}P\{X\geqslant n\}.$$

它提供了一个计算数学期望的公式.

如果不用这一技巧,可以这样证明:

$$\begin{aligned}\sum_{n=1}^{\infty}P\{X\geqslant n\}&=\sum_{n=1}^{\infty}\sum_{k=n}^{\infty}P\{X=k\}\\&=P\{X=1\}+P\{X=2\}+P\{X=3\}+\cdots\\&\qquad\qquad+P\{X=2\}+P\{X=3\}+\cdots\\&\qquad\qquad\qquad\qquad+P\{X=3\}+\cdots\\&\qquad\qquad\qquad\qquad\qquad\qquad+\cdots\\&=P\{X=1\}+2P\{X=2\}+3P\{X=3\}+\cdots\\&=\sum_{k=0}^{\infty}kP\{X=k\}=E(X).\end{aligned}$$

显然要麻烦得多.

例 29 设 X 是取非负整数值的随机变量. 证明:

$$D(X)=2\sum_{n=1}^{\infty}nP\{X\geqslant n\}-E(X)[E(X)+1].$$

证 如同例 29,利用级数交换和号的技巧,有

$$\begin{aligned}\sum_{n=1}^{\infty}(2n-1)P\{X\geqslant n\}&=\sum_{n=1}^{\infty}(2n-1)\sum_{k=n}^{\infty}P\{X=k\}\\&=\sum_{k=1}^{\infty}P\{X=k\}\sum_{n=1}^{\infty}(2n-1)\\&=\sum_{k=1}^{\infty}k^2P\{X=k\}=E(X^2),\end{aligned}$$

利用例29的结果，得

$$\sum_{n=1}^{\infty}(2n-1)P\{X\geqslant n\}=2\sum_{n=1}^{\infty}nP\{X\geqslant n\}-\sum_{n=1}^{\infty}P\{X\geqslant n\}$$
$$=2\sum_{n=1}^{\infty}nP\{X\geqslant n\}=E(X).$$

所以 $$D(X)=E(X^2)-[E(X)^2]$$
$$=2\sum_{n=1}^{\infty}nP\{X\geqslant n\}-E(X)-[E(X)]^2$$
$$=2\sum_{n=1}^{\infty}nP\{X\geqslant n\}-E(X)[E(X)+1].$$

例30 设随机变量$X\sim e(\lambda)$，证明：使$E[|X-C|]$取得最小值的常数$C=\frac{1}{\lambda}\ln 2$，且C满足

$$P\{X\leqslant C\}=P\{X\geqslant C\}=1/2.$$

证 作函数$\varphi(C)=E[|X-C|]$，因为

$$f(x)=\begin{cases}\lambda e^{-\lambda x}, & x\geqslant 0,\\ 0, & 其它,\end{cases}$$

所以 $$E[|X-C|]=\int_0^C\lambda(C-x)e^{-\lambda x}dx+\int_C^{+\infty}\lambda(x-C)e^{-\lambda x}dx$$
$$=C+\frac{1}{\lambda}\left(2e^{-\lambda C}-\frac{1}{\lambda}\right).$$

求导得 $$\varphi'(C)=1-2e^{-\lambda C},\quad \varphi''(C)=2\lambda e^{-\lambda C}.$$

令$\varphi'(C)=0\Longrightarrow C=\frac{1}{\lambda}\ln 2$为驻点，又$\varphi''(C)>0$，故在$C=\frac{1}{\lambda}\ln 2$时，$\varphi(C)=E[|X-C|]$为极小值（即最小值）点. 求出最小值为$\frac{1}{\lambda}\ln 2$.

又 $$P\{X\leqslant C\}=\int_0^C\lambda e^{-\lambda x}dx=1-e^{-\lambda C},$$

当$C=\frac{1}{\lambda}\ln 2$时，$P\{X\leqslant C\}=\frac{1}{2}$. 从而

$$P\{X\leqslant C\}=P\{X\geqslant C\}=1/2.$$

一般，把使$P\{X\geqslant C\}=P\{X\leqslant C\}$的点$x=C$称为$X$的中位数.

一般可证，使 $E[|X-C|]$ 达到最小值的 C 即为中位数.

例 31 设随机变量 X 的密度函数为 $f(x)$，若对于常数 C，有

$$f(C+x)=f(C-x), \quad x>0$$

且 $E(X)$ 存在，证明：$E(X)=C$.

证 要用积分技巧来证. 因为

$$E(X)=\int_{-\infty}^{+\infty}xf(x)\mathrm{d}x \xlongequal{\text{令 } t=x-C} \int_{-\infty}^{+\infty}(C+t)f(C+t)\mathrm{d}t$$

$$=\int_{-\infty}^{+\infty}Cf(C+t)\mathrm{d}t+\int_{-\infty}^{+\infty}tf(C+t)\mathrm{d}t,$$

而 $$\int_{-\infty}^{+\infty}Cf(C+t)\mathrm{d}t \xlongequal{\text{令 } C+t=u} \int_{-\infty}^{+\infty}f(u)\mathrm{d}u=C\times 1=C,$$

$$\int_{-\infty}^{0}tf(C+t)\mathrm{d}t=\int_{-\infty}^{0}tf(C-t)\mathrm{d}t \quad (\text{由题设})$$

$$\xlongequal{\text{令 } u=-t} \int_{-\infty}^{0}uf(C+u)\mathrm{d}u=-\int_{0}^{+\infty}uf(C+u)\mathrm{d}u,$$

所以 $$\int_{-\infty}^{+\infty}tf(C+t)=\int_{-\infty}^{0}tf(C+t)\mathrm{d}t+\int_{0}^{+\infty}tf(C+t)\mathrm{d}t=0,$$

$$E(X)=C+0=C.$$

例32 设随机变量 X 有 $E(X)=\mu, D(X)=\sigma^2$，且 $Y=g(X)$，证明：

$$E(Y)\approx g(\mu)+\frac{1}{2}g''(\mu)\sigma^2, \quad D(Y)\approx[g'(\mu)]^2\sigma^2,$$

其中 $g(x)$ 在 $x=\mu$ 处的二阶导数存在.

证 将 $Y=g(X)$ 在 $X=\mu$ 处展开为二阶泰勒公式

$$Y=g(\mu)+g'(\mu)(X-\mu)+\frac{1}{2}g''(\mu)(X-\mu)^2+R_2(X),$$

舍去余项 $R_2(X)$，对上式两边取期望，得

$$E(Y)\approx g(\mu)+\frac{1}{2}g''(\mu)E[(X-\mu)^2]=g(\mu)+\frac{1}{2}g''(\mu)\sigma^2.$$

将 $Y=g(X)$ 在 $X=\mu$ 处展开为一阶泰勒公式

$$Y=g(\mu)+g'(\mu)(X-\mu)+R_1(X),$$

舍去余项 $R_1(X)$，对上式两边取方差，得

$$D(Y)\approx[g'(\mu)]^2D(X)=[g'(\mu)]^2\sigma^2.$$

例 33 证明：若随机变量 X 有 $E(X^2)$ 存在，则 $E(X)$ 也存在.

证 因为对随机变量 X，利用不等式性质，由

$$(|X|+1)^2\geqslant 0,$$

得
$$0\leqslant|X|\leqslant\frac{1}{2}(X^2+1),$$

所以
$$E(|X|)\leqslant\frac{1}{2}E(X^2+1)=\frac{1}{2}E(X^2)+1<+\infty,$$

故 $E(X)$ 存在.

例 34 设 X 为随机变量，C 为任意常数，且 $C\neq E(X)$，证明：

$$D(X)<E[(X-C)^2].$$

证 用组合的技巧，得

$$D(X)=E\{[X-E(X)]^2\}=E\{[(X-C)-E(X)-C]^2\}$$
$$=E[(X-C)^2]-[E(X)-C]^2,$$

移项
$$D(X)+[E(X)-C]^2=E[(X-C)^2],$$

所以
$$D(X)<E[(X-C)^2].$$

例 35 设随机变量 X 与 Y 相互独立且同分布，令 $U=X+Y$，$V=X-Y$，证明：U,V 必然不相关.

证 只需求 $\rho_{UV}=0$，可转化为求 $\mathrm{cov}(U,V)=0$.

$$\mathrm{cov}(U,V)=E(UV)-E(U)E(V)$$
$$=E[(X+Y)(X-Y)]-E(X+Y)E(X-Y)$$
$$=E(X^2)-E(Y^2)-[E(X)]^2+[E(Y)]^2$$
$$=D(X)-D(Y)=0,$$

故 $\rho_{UV}=0$，U 与 V 不相关.

例 36 设随机变量 X 与 Y 相互独立且同分布，令 $U=\alpha X+\beta Y$，$V=\alpha X-\beta Y$，其中 α,β 为常数，证明：

$$\rho_{UV}=\frac{\alpha^2-\beta^2}{\alpha^2+\beta^2}.$$

证 因为 X 与 Y 相互独立且同分布，所以

$$E(X)=E(Y)=\mu,\quad D(X)=D(Y)=\sigma^2,$$

于是 $E(U)=E(\alpha X+\beta Y)=(\alpha+\beta)\mu$，$E(V)=(\alpha-\beta)\mu$，

$$D(U)=D(\alpha X+\beta Y)=(\alpha^2+\beta^2)\sigma^2,\quad D(V)=(\alpha^2+\beta^2)\sigma^2,$$

从而

$$\begin{aligned}\operatorname{cov}(U,V)&=E[(\alpha X+\beta Y)(\alpha X-\beta Y)]-E(\alpha X+\beta Y)E(\alpha X-\beta Y)\\&=\alpha^2D(X)-\beta^2D(Y)=(\alpha^2-\beta^2)\sigma^2,\end{aligned}$$

得

$$\rho_{UV}=\frac{(\alpha^2-\beta^2)\sigma^2}{(\alpha^2+\beta^2)\sigma^2}=\frac{\alpha^2-\beta^2}{\alpha^2+\beta^2}.$$

显然，当 $|\alpha|=|\beta|$ 时，U 与 V 不相关. 例35为本例的特例.

例 37 设随机变量 X 和 Y 相互独立，证明：

$$D(XY)=D(X)D(Y)+[E(X)]^2D(Y)+[E(Y^2)]D(X).$$

证 因为 X 与 Y 相互独立，所以

$$E(XY)=E(X)E(Y),\quad E(X^2Y^2)=E(X^2)E(Y^2),$$

$$\begin{aligned}D(XY)&=E(X^2Y^2)-[E(XY)]^2\\&=E(X^2)E(Y^2)-[E(X)]^2[E(Y)]^2\\&=\{D(X)+[E(X)]^2\}E(Y^2)-[E(X)]^2[E(Y)]^2\\&=D(X)E(Y^2)+[E(X)]^2E(Y^2)-[E(X)]^2[E(Y)]^2\\&=D(X)E(Y^2)+[E(X)]^2\{E(Y^2)-[E(Y)]^2\}\\&=D(X)E(Y^2)+[E(X)]^2D(Y)\\&=D(X)D(Y)+D(X)[E(Y)]^2+[E(X)]^2D(Y).\end{aligned}$$

其中多次利用了随机变量方差与数学期望的关系式

$$D(X)=E(X^2)-[E(X)]^2.$$

例38 设 X 和 Y 为两个随机变量，$E(X)=E(Y)=0$，$D(X)=D(Y)=1$，$\rho_{XY}=\operatorname{cov}(X,Y)$，证明：

$$E[\max(X^2,Y^2)]\leqslant 1+\sqrt{1-\rho^2}.$$

证 与上例相同，利用数学期望与方差的关系式，特别利用了 $E(X)=0$，$E(Y)=0$，$E(X+Y)=0$.

$$\begin{aligned}&E[\max(X^2,Y^2)]\\&=E\left[\frac{1}{2}(X^2+Y^2+|X^2-Y^2|)\right]\end{aligned}$$

$$=\frac{1}{2}[E(X^2)+E(Y^2)+E|X^2-Y^2|]$$

$$=\frac{1}{2}[D(X)+D(Y)+E(|X+Y||X-Y|)]$$

$$\leqslant\frac{1}{2}\left[D(X)+D(Y)+\sqrt{D(X)+D(Y)+2\mathrm{cov}(X,Y)}\right.$$

$$\left.\cdot\sqrt{D(X)+D(Y)-2\mathrm{cov}(X,Y)}\right]$$

$$=\frac{1}{2}\left[D(X)+D(Y)+\sqrt{D(X)+D(Y)+2\rho\sqrt{D(X)D(Y)}}\right.$$

$$\left.\cdot\sqrt{D(X)+D(Y)-2\rho\sqrt{D(X)D(Y)}}\right]$$

$$\xlongequal{\text{(代入)}}\frac{1}{2}\left[1+1+\sqrt{(1+1+2\rho)(1+1-2\rho)}\right]$$

$$=1+\sqrt{1-\rho^2},$$

即 $$E[\max(X^2,Y^2)]\leqslant 1+\sqrt{1-\rho^2}\ (\text{式中},\rho\text{ 即 }\rho_{XY}).$$

例 39 设 $X_1,X_2,\cdots,X_n$ 是相互独立的随机变量，$D(X_i)=\sigma_i^2$. α_i $(i=1,2,\cdots,n)$为常数，且$\sum\limits_{i=1}^n\alpha_i=1$. 证明：使 $D\left[\sum\limits_{i=1}^n\alpha_iX_i\right]$达到最小值的 α_i 为

$$\alpha_i=\frac{1}{\sigma_i^2}\Big/\sum_{k=1}^n\frac{1}{\sigma_k^2},\quad i=1,2,\cdots,n.$$

证 由 $X_1,X_2,\cdots,X_n$ 的独立性知

$$D\left(\sum_{i=1}^n\alpha_iX_i\right)=\sum_{i=1}^n\alpha_i^2D(X_i)=f(\alpha_1,\alpha_2,\cdots,\alpha_n).$$

现求函数 $f(\alpha_1,\alpha_2,\cdots,\alpha_n)$在约束条件$\sum\limits_{i=1}^n\alpha_i=1$ 下的条件极值. 建立拉格朗日函数

$$L=\sum_{i=1}^n\alpha_i^2\sigma_i^2+\lambda\left(\sum_{i=1}^n\alpha_i-1\right),$$

由 $$\begin{cases}\dfrac{\partial L}{\partial\alpha_i}=2\alpha_i\sigma_i^2+\lambda=0,\\ \sum\limits_{i=1}^n\alpha_i-1=0,\end{cases}$$

得 $$\alpha_i=-\frac{\lambda}{2\sigma_i^2},\quad \lambda=-1\Big/\sum_{i=1}^{n}\frac{1}{\sigma_i^2}.$$

所以，当 $\alpha_i=\frac{1}{\sigma_i^2}\Big/\sum_{k=1}^{n}\frac{1}{\sigma_k^2}\ (i=1,2,\cdots,n)$时，$D\left(\sum_{i=1}^{n}\alpha_iX_i\right)$达到最小值. 这时，方差为

$$D\left(\sum_{i=1}^{n}\alpha_iX_i\right)=1\Big/\sum_{i=1}^{n}\frac{1}{\sigma_i^2}.$$

例 40 设 A,B 是两随机事件，定义

$$X=\begin{cases}1, & A\text{ 发生},\\ 0, & A\text{ 不发生},\end{cases}\qquad Y=\begin{cases}1, & B\text{ 发生},\\ 0, & B\text{ 不发生},\end{cases}$$

证明：$\rho_{XY}=0\Longleftrightarrow X$ 与 Y 相互独立.

证 利用事件独立性证. 写出 X,Y 和 XY 的分布律

X	0	1
p	$P(A)$	$P(\overline{A})$

Y	0	1
p	$P(B)$	$P(\overline{B})$

XY	0	1
p	$1-P(AB)$	$P(AB)$

则 $E(X)=P(A),\quad E(Y)=P(B),\quad E(XY)=P(AB).$

又 $$\rho_{XY}=0\Longrightarrow \mathrm{cov}(X,Y)=0,$$

故 $$E(XY)=E(X)E(Y),$$

即 $$P(AB)=P(A)P(B),$$

知事件 A 与 B 相互独立，从而 A 与 $\overline{B}$，$\overline{A}$ 与 B，$\overline{A}$ 与 $\overline{B}$ 也相互独立，于是

$$P\{X=1,Y=1\}=P(AB)=P(A)P(B)=P\{X=1\}P\{Y=1\},$$
$$P\{X=1,Y=0\}=P(A\overline{B})=P(A)P(\overline{B})=P\{X=1\}P\{Y=0\},$$
$$P\{X=0,Y=1\}=P(\overline{A}B)=P(\overline{A})P(B)=P\{X=0\}P\{Y=1\},$$
$$P\{X=0,Y=0\}=P(\overline{A}\overline{B})=P(\overline{A})P(\overline{B})=P\{X=0\}P\{Y=0\}.$$

所以 X 与 Y 相互独立.

例 41 设 A,B 是两随机事件，定义

$$X=\begin{cases} 1, & A\text{ 发生}, \\ -1, & A\text{ 不发生}, \end{cases} \quad Y=\begin{cases} 1, & B\text{ 发生}, \\ -1, & B\text{ 不发生}. \end{cases}$$

证明：$\rho_{XY}=0 \Longleftrightarrow A$ 与 B 相互独立.

证 记

$$P(A)=p_1, \quad P(B)=p_2, \quad P(AB)=p_{12},$$

由数学期望的定义知

$$E(X)=1\times P(A)-1\times P(\overline{A})=p_1-(1-p_1)=2p_1-1,$$

$$E(Y)=1\times P(B)-1\times P(\overline{B})=2p_2-1.$$

因为 XY 只有两个可取值 $-1,1$，故

$$P\{XY=1\}=P(AB)+P(\overline{A}\overline{B})=2p_{12}-p_1-p_2+1,$$

$$P\{XY=-1\}=1-P\{XY=1\}=p_1+p_2-2p_{12},$$

所以 $$E(XY)=1\times P\{XY=1\}-1\times P\{XY=-1\}$$
$$=4p_{12}-2p_1-2p_2+1.$$

于是 $$\mathrm{cov}(X,Y)=E(XY)-E(X)E(Y)=4p_{12}-4p_1p_2.$$

从而知，要 $\rho_{XY}=0$，必有 $p_{12}=p_1p_2$，即 A 与 B 相互独立.

例41 与例42 十分相似，但两题证明方法不同，结论也不同. 读者可以尝试将两例的方法调换进行证明，看是否能证出.

例 42 设随机变量 $X_1,X_2,\cdots,X_{2n}$ 有相同的均值 μ 和方差 σ^2，且当 $i\neq j$ 时，$\rho_{X_iX_j}=\rho$. 令 $U=\sum_{i=1}^{n}X_i, V=\sum_{k=1}^{n}X_{n+k}$，求 ρ_{UV}.

证 $$E(U)=E\Big(\sum_{i=1}^{n}X_i\Big)=n\mu, \quad E(V)=n\mu,$$

$$\mathrm{cov}(X_i,X_j)=\rho_{X_iX_j}\sqrt{D(X_i)}\sqrt{D(X_j)}=\rho\sigma^2,$$

所以 $$D(U)=D\Big(\sum_{i=1}^{n}X_i\Big)=\sum_{i=1}^{n}D(X_i)+\sum_{i,j=1;i\neq j}^{n}\mathrm{cov}(X_i,X_j)$$
$$=n\sigma^2+(n^2-n)\rho\sigma^2,$$

$$D(V)=n\sigma^2+(n^2-n)\rho\sigma^2.$$

又 $$E(UV)=E\Big(\sum_{i=1}^{n}X_i\cdot\sum_{k=1}^{n}X_{n+k}\Big)=\sum_{i,k=1}^{n}E(X_iX_{n+k}),$$

由 $\rho_{X_iX_{n+k}}=[E(X_iX_{n+k})-E(X_i)E(X_{n+k})]\Big/\sqrt{D(X_i)D(X_{n+k})}$

可得 $$E(X_iX_{n+k})=\rho\sqrt{D(X_i)D(X_{n+k})}+E(X_i)E(X_{n+k})$$
$$=\rho\sigma^2+\mu^2,$$

所以

$$\mathrm{cov}(U,V)=E(U,V)-E(U)E(V)$$
$$=\sum_{i,k=1}^{n}E(X_iX_{n+k})-n\mu^2$$
$$=n^2(\rho\sigma^2+\mu^2)-n^2\mu^2=n^2\rho\sigma^2.$$

从而 $$\rho_{UV}=\frac{\mathrm{cov}(U,V)}{\sqrt{D(U)}\sqrt{D(V)}}=\frac{n^2\rho\sigma^2}{n\sigma^2+(n^2-n)\rho\sigma^2}$$
$$=\frac{n\rho}{1+(n-1)\rho}.$$

例 43 设 $g(x)$ 是正值不减函数，X 是连续型随机变量，且 $g(X)$ 的数学期望存在. 证明：

$$P\{X\geqslant a\}\leqslant E[g(x)]/g(a).$$

证 设 X 的概率密度为 $f(x)$，于是

$$P\{X\geqslant a\}=P\{g(x)\geqslant g(a)\}=\int_{g(x)\geqslant g(x)}f(x)\mathrm{d}x$$
$$\leqslant\int_{g(x)\geqslant g(a)}\frac{g(x)}{g(a)}f(x)\mathrm{d}x$$
$$\leqslant\frac{1}{g(a)}\int_{-\infty}^{+\infty}g(x)f(x)\mathrm{d}x$$
$$=E[g(x)]/g(a).$$

例 44 对两个随机变量 X,Y，若 $E(X^2),E(Y^2)$ 存在，证明：

$$[E(XY)]^2\leqslant E(X^2)E(Y^2)\quad(\text{柯西-许瓦兹不等式}).$$

解 引入实变量 t 的函数

$$g(t)=E[(X+Yt)^2]=E(X^2)+2tE(XY)+t^2E(Y^2).$$

因为，对任何实数 t，$g(t)\geqslant 0$，故

$$\Delta=4[E(XY)]^2-4E(X^2)E(Y^2)\leqslant 0,$$

即 $$[E(XY)]^2\leqslant E(X^2)E(Y^2).$$

硕士研究生入学试题分析

一、本章考试要求

1. 理解随机变量数字特征(数学期望、方差、标准差、协方差、矩相关系数)的概念,会运用数字特征的基本性质计算具体分布的数字特征,并掌握常用分布的数字特征.

2. 会根据随机变量 X 的分布求其函数 $g(X)$ 的数学期望 $E[g(X)]$,会根据 X 和 Y 的联合概率分布求其函数 $g(X,Y)$ 的数学期望 $E[g(X,Y)]$.

二、本章重点内容

考试中,很重要的一类题型就是综合题.它与各章节比较单一的习题不同,往往将几章的知识联系在一起,解题的步骤也比较复杂.读者要学会通过审题由已知可以得到哪些结果,再由此结果而逐步求得我们所需结果.其它考点还有:计算数字特征、通过数字特征确定参数、通过数字特征讨论相关性与独立性、解关于数字特征的应用题.

(一) 数学期望与方差

1. 设随机变量 X 和 Y 都服从正态分布,且它们不相关,则(　　).

(A) X 与 Y 一定相互独立;　(B) (X,Y) 服从二维正态分布;

(C) X 与 Y 未必相互独立;　(D) $X+Y$ 服从一维正态分布.

(2003 年四)

解　选(C). X 与 Y 不相关, X 与 Y 未必相互独立.

首先,两个正态随机变量的联合分布不一定是正态的.如 $X\sim N(0,1)$, $Y\sim N(0,1)$,而

$$f(x,y)=\frac{1}{2\pi}\exp\left[-\frac{1}{2}(x^2+y^2)\right]\cdot(1+\sin x\sin y)$$

就不是正态分布，故(B)不成立. 仅当(X,Y)服从二维正态分布时，X与Y相互独立，故(A)也不成立. 仅当X与Y相互独立时，$X+Y$服从一维正态分布，故(D)也不成立.

2. 设随机变量$X_1,X_2,\cdots,X_n$ $(n>1)$相互独立且同分布，且其方差为$\sigma^2>0$，令$Y=\frac{1}{n}\sum_{i=1}^{n}X_i$，则(　　).

(A) $\mathrm{cov}(X_1,Y)=\sigma^2/n$;　　(B) $\mathrm{cov}(X_1,Y)=\sigma^2$;

(C) $D(X_1+Y)=\frac{n+2}{n}\sigma^2$;　　(D) $D(X_1-Y)=\frac{n+1}{n}\sigma^2$.

(2004 年一、四)

解 选(A). 由$\mathrm{cov}(X_1,Y)=\frac{1}{n}\mathrm{cov}\left(X_1,\sum_{i=1}^{n}X_i\right)$而得.

3. 设随机变量X和Y的相关系数为0.9，若$Z=X-0.4$，则Y与Z的相关系数为________. (2003 年三)

解 因为$D(Z)=D(X)$，　$\mathrm{cov}(Z,Y)=\mathrm{cov}(X,Y)$，所以

$$\rho_{YZ}=\rho_{XY}=0.9.$$

4. 设随机变量X服从参数为λ的指数分布，则

$$P\{X>\sqrt{D(X)}\}=________.$$

(2004 年一、三、四)

解 由题设知，X的概率密度为

$$f(x)=\begin{cases}\lambda e^{-\lambda x}, & x\geqslant 0,\\ 0, & x<0,\end{cases}$$

$$E(X)=1/\lambda,\quad D(X)=1/\lambda^2,$$

故
$$P\{X>\sqrt{D(X)}\}=P\{X>1/\lambda\}=\int_{1/\lambda}^{\infty}\lambda e^{-\lambda x}dx$$
$$=-\int_{1/\lambda}^{\infty}e^{-\lambda x}d(-\lambda x)=-e^{-\lambda x}\Big|_{1/\lambda}^{\infty}=\frac{1}{e}.$$

5. 设X是一随机变量，$E(X)=\mu,D(X)=\sigma^2$ $(\mu,\sigma>0$，常数)，则对任意常数C，必有(　　).

(A) $E(X-C)^2=E(X^2)-C^2$;

(B) $E(X-C)^2=E(X-\mu)^2$;

(C) $E(X-C)^2<E(X-\mu)^2$;

(D) $E(X-C)^2\geqslant E(X-\mu)^2$. (1997 年四)

解 选(D). 设 $f(C)=E(X-C)^2$,有

$$E(X-C)^2=E(X^2+C^2-2CX)=E(X^2)+C^2-2CE(X).$$

令 $f'(C)=2C-2E(X)=0$,得驻点 $C=E(X)$. 又 $f''(C)=2>0$,故当 $C=E(X)=\mu$ 时有最小值,即 $E(X-\mu)^2\leqslant E(X-C)^2$.

6. 设两个相互独立的随机变量 X 和 Y 的方差分别为4 和2,则随机变量 $3X-2Y$ 的方差是(   ).

(A) 8; (B) 16; (C) 28; (D) 44. (1997 年一)

解 选(D),因为

$$D(3X-2Y)=9D(X)+4D(Y)=9\times 4+4\times 2=44.$$

7. 对于任意两个随机变量 X 与 Y,若 $E(XY)=E(X)E(Y)$,则(   ).

(A) $D(XY)=D(X)D(Y)$;

(B) $D(X+Y)=D(X)+D(Y)$;

(C) X 和 Y 相互独立;

(D) X 和 Y 不相互独立. (1991 年四)

解 选(B),因为由相关函数性质有

$$E(XY)=E(X)E(Y)\sim\rho_{XY}=0\sim D(X+Y)=D(X)+D(Y).$$

8. 设随机变量 X_{ij} $(i,j=1,2,\cdots,n,n\geqslant 2)$ 相互独立且同分布,$E(X_{ij})=2$,则行列式

$$Y=\begin{vmatrix} X_{11} & X_{12} & \cdots & X_{1n} \\ X_{21} & X_{22} & \cdots & X_{2n} \\ \vdots & \vdots & & \vdots \\ X_{n1} & X_{n2} & \cdots & X_{nn} \end{vmatrix}$$

的数学期望 $E(Y)=$________. (1999 年三)

解 由行列式性质可得

$$E(Y)=E\begin{vmatrix} X_{11} & X_{12} & \cdots & X_{1n} \\ X_{21} & X_{22} & \cdots & X_{2n} \\ \vdots & \vdots & & \vdots \\ X_{n1} & X_{n2} & \cdots & X_{nn} \end{vmatrix}$$

$$=\begin{vmatrix} E(X_{11}) & E(X_{12}) & \cdots & E(X_{1n}) \\ E(X_{21}) & E(X_{22}) & \cdots & E(X_{2n}) \\ \vdots & \vdots & & \vdots \\ E(X_{n1}) & E(X_{n2}) & \cdots & E(X_{nn}) \end{vmatrix}$$

$$=\begin{vmatrix} 2 & 2 & \cdots & 2 \\ 2 & 2 & \cdots & 2 \\ \vdots & \vdots & & \vdots \\ 2 & 2 & \cdots & 2 \end{vmatrix}=0.$$

9. 设 X 是一个随机变量,其概率密度为

$$f(x)=\begin{cases} 1+x, & -1\leqslant x<0, \\ 1-x, & 0\leqslant x\leqslant 1, \\ 0, & \text{其它}, \end{cases}$$

则方差 $D(X)=$________. (1995 年五)

解 $$E(X)=\int_{-1}^{0}(1+x)x\mathrm{d}x+\int_{0}^{1}(1-x)\mathrm{d}x=0,$$

$$E(X^2)=\int_{-1}^{0}(1+x)^2\mathrm{d}x\int_{0}^{1}(1-x)x^2\mathrm{d}x=1/6,$$

所以 $$D(X)=E(X^2)-[E(X)]^2=1/6,$$

10. 设 ξ 和 η 是两个相互独立且服从正态分布 $N(0,(1/\sqrt{2})^2)$ 的随机变量,则随机变量 $|\xi-\eta|$ 的数学期望 $E|\xi-\eta|=$________. (1996 年一)

解 令 $Z=\xi-\eta$,则 $Z\sim N(0,1)$,于是

$$E|\xi-\eta|=E|Z|=\int_{-\infty}^{+\infty}|z|\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-z^2/2}\mathrm{d}z=\sqrt{2/\pi}.$$

11. 已知离散型随机变量 X 服从参数为 2 的泊松分布,即 $P\{X=k\}=2^k\mathrm{e}^{-2}/k!$,则随机变量 $Y=3X-2$ 的数学期望 $E(Y)=$

__________. (1990 年一)

解 $E(X)=\lambda=2$,所以

$$E(Y)=E(3X-2)=3E(X)-2=3\times2-2=4.$$

12. 已知甲、乙两箱中装有同种产品,其中甲箱中装有3件合格品、3件次品,乙箱中仅装有3件合格品,从甲箱中任取3件产品放入乙箱后,求:

(1) 乙箱中次品件数X的数学期望;

(2) 从乙箱中任取一件产品是次品的概率.

(2003 年一)

解一 (1) 设 $X_i=\begin{cases}0, & \text{从甲箱中取出第}i\text{件产品是合格品},\\ 1, & \text{从甲箱中取出第}i\text{件产品是次品},\end{cases}$

则X_i的分布律为

X_i	0	1
p_k	1/2	1/2

$i=1,2,3$,且$E(X_i)=1/2$. 由于$X=\sum\limits_{i=1}^{3}X_i$,故

$$E(X)=E\Big(\sum_{i=1}^{3}X_i\Big)=\sum_{i=1}^{3}E(X_i)=\frac{3}{2}.$$

(2) 以A记事件{从乙箱中任取一件是次品},则$\{X=0\}$,$\{X=1\}$,$\{X=2\}$,$\{X=3\}$构成完备事件组,由全概率公式

$$P(A)=\sum_{k=0}^{3}P\{X=k\}P\{A\,|\,X=k\}=\sum_{k=0}^{3}P\{X=k\}\cdot\frac{k}{6}$$

$$=\frac{1}{6}\sum_{k=0}^{3}kP\{X=k\}=\frac{1}{6}E(X)=\frac{1}{6}\times\frac{3}{2}=\frac{1}{4}.$$

解二 (1) X的可能取值为0,1,2,3,X的分布律为

$$P\{X=k\}=C_3^kC_3^{3-k}/C_6^3,\quad k=0,1,2,3,$$

即

X	0	1	2	3
p_k	1/20	9/20	9/20	1/20

故 $$E(X)=kP\{X=k\}=3/2.$$

13. 对于任意两事件 A 和 B,$0<P(A)<1,0<P(B)<1$,

$$\rho=\frac{P(AB)-P(A)P(B)}{\sqrt{P(A)P(B)P(\overline{A})P(\overline{B})}}$$

称为事件 A 和 B 的相关系数.

(1) 证明事件 A 与 B 相互独立的充分必要条件是其相关系数等于零;

(2) 利用随机变量相关系数的基本性质,证明 $|\rho|\leqslant 1$.

(2003 年四)

证 (1) 由题给定义知,当且仅当 $P(AB)=P(A)P(B)$ 时 $\rho=0$,即 A 与 B 相互独立,故 $\rho=0$ 是 A 与 B 相互独立的充分必要条件.

(2) 为证明 $|\rho|\leqslant 1$,设随机变量

$$X=\begin{cases}1, & A\text{ 发生},\\ 0, & A\text{ 不发生},\end{cases}\qquad Y=\begin{cases}1, & B\text{ 发生},\\ 0, & B\text{ 不发生},\end{cases}$$

则 X 与 Y 都服从 0-1 分布,有

$$X\sim\begin{pmatrix}0 & 1\\ 1-P(A) & P(A)\end{pmatrix},\quad Y\sim\begin{pmatrix}0 & 1\\ 1-P(B) & P(B)\end{pmatrix}.$$

于是 $$E(X)=P(A),\quad E(Y)=P(B),$$
$$D(X)=P(A)P(\overline{A}),\quad D(Y)=P(B)P(\overline{B}),$$
$$\operatorname{cov}(X,Y)=P(AB)-P(A)P(B).$$

从而,事件 A 与 B 的相关系数即随机变量 X 和 Y 的相关系数. 由两随机变量相关系数的基本性质 $|\rho_{XY}|\leqslant 1$ 知,$|\rho|\leqslant 1$.

14. 设随机变量 X 和 Y 同分布,X 的概率密度为

$$f(x)=\begin{cases}\dfrac{3}{8}x^2, & 0<x<2,\\ 0, & \text{其它}.\end{cases}$$

(1) 已知事件 $A=\{X>a\}$ 和 $B=\{Y>a\}$ 独立,且 $P\{A\cup B\}=3/4$,求常数 a;

(2) 求 $1/X^2$ 的数学期望. (1993 年四)

解 (1) 因为$P(A)=P(B)$，$P(AB)=P(A)P(B)$，

$$P(A+B)=P(A)+P(B)-P(AB)$$
$$=2P(A)-[P(A)]^2=3/4,$$

所以
$$P(A)=1/2.$$

$$P\{X>a\}=\int_a^{+\infty}f(x)\mathrm{d}x=\frac{3}{8}\int_a^2 x^2\mathrm{d}x=1-\frac{a^3}{8}.$$

由$P\{X>a\}=P(A)$，得$a=\sqrt[3]{4}$.

(2) $E(1/X^2)=\int_{-\infty}^{+\infty}\frac{1}{x^2}f(x)\mathrm{d}x=\frac{3}{8}\int_0^2\mathrm{d}x=\frac{3}{4}$.

15. 设随机变量X和Y独立，都在区间$[1,3]$上服从均匀分布，引进事件$A=\{X\leqslant a\}$，$B=\{Y>a\}$.

(1) 已知$P\{A\cup B\}=7/9$，求常数a；

(2) 求$1/X$的数学期望. (1993年五)

解 (1) 设$P(A)=p$，则

$$P(\overline{B})=P(A)=p,\quad P(B)=1-p.$$

由条件知

$$P(A+B)=P(A)+P(B)-P(A)P(B)=p^2-p+1=7/9,$$

解得
$$p_1=1/3,\quad p_2=2/3.$$

从而
$$a_1=1+2/3=5/3,\quad a_2=1-4/3=7/3.$$

(2) $E(1/X)=\int_{-\infty}^{+\infty}\frac{1}{x}f(x)\mathrm{d}x=\frac{1}{2}\int_1^3 1/x\mathrm{d}x=\frac{1}{2}\ln 3$.

16. 从学校乘汽车到火车站途中有三个交通岗，假设在各个交通岗遇到红灯的事件是相互独立的，且概率都是2/5. 设X为途中遇到红灯的次数，求随机变量X的分布律、分布函数和数学期望. (1997年一)

解 X服从二项分布$B(3,2/5)$，X可能取值为0,1,2,3，故

$$P\{X=0\}=(1-2/5)^3=27/125,$$
$$P\{X=1\}=C_3^1\times 2/5\times(1-2/5)^2=54/125,$$
$$P\{X=2\}=C_3^2\times(2/5)^2\times(1-2/5)=36/125,$$
$$P\{X=3\}=(2/5)^3=8/125.$$

即 X 的分布函数为

$$F(x)=\begin{cases} 0, & x<0, \\ 27/125, & 0\leqslant x<1, \\ 81/125, & 1\leqslant x<2, \\ 117/125, & 2\leqslant x<3, \\ 1, & x\geqslant 3. \end{cases}$$

分布律为

X	0	1	2	3
p_k	27/125	54/125	36/125	8/125

数学期望为 $$E(X)=np=3\times 2/5=6/5.$$

17. 设 ξ 和 η 是相互独立且服从同一分布的两个随机变量，已知 ξ 的分布律为 $P\{\xi=i\}=1/3, i=1,2,3$. 又设

$$X=\max\{\xi,\eta\}, \quad Y=\min\{\xi,\eta\}.$$

(1) 写出二维随机变量 (X,Y) 的分布律；

(2) 求随机变量 X 的数学期望 $E(X)$. （1996 年一）

解 (1) 因为 $P\{\xi=i\}=1/3,\ i=1,2,3$，所以

$$P\{X=1\}=P\{\xi=1,\eta=1\}=1/3\times 1/3=1/9,$$

$$P\{X=2\}=P\{\xi=2,\eta=1\}+P\{\xi=2,\eta=2\}+P\{\xi=1,\eta=2\}$$
$$=1/3,$$

$$P\{X=3\}=P\{\xi=3,\eta=1\}+P\{\xi=1,\eta=3\}+P\{\xi=3,\eta=2\}$$
$$+P\{\xi=2,\eta=3\}+P\{\xi=3,\eta=3\}$$
$$=5/9,$$

$$P\{Y=1\}=P\{\xi=1,\eta=1\}+P\{\xi=1,\eta=2\}+P\{\xi=2,\eta=1\}$$
$$+P\{\xi=1,\eta=3\}+P\{\xi=3,\eta=1\}$$
$$=5/9,$$

$$P\{Y=2\}=P\{\xi=2,\eta=3\}+P\{\xi=3,\eta=2\}+P\{\xi=2,\eta=2\}$$
$$=1/3,$$

$$P\{Y=3\}=P\{\xi=3,\eta=3\}=1/9.$$

分布律为

$Y \backslash X$	1	2	3	$p_{\cdot j}$
1	1/9	2/9	2/9	5/9
2	0	1/9	2/9	1/3
3	0	0	1/9	1/9
$p_{i\cdot}$	1/9	1/3	5/9	

(2) $E(X)=1\times1/9+2\times1/3+3\times5/9=22/9$.

18. 设随机变量 Y 服从参数为 1 的指数分布，随机变量

$$X_k=\begin{cases}0, & Y\leqslant k,\\ 1, & Y>k,\end{cases}\quad k=1,2,$$

求：(1) X_1 和 X_2 的联合分布； (2) $E(X_1+X_2)$.　　(1997 年四)

解　(1) 因为 $F(y)=\begin{cases}1-e^{-y}, & y<0,\\ 0, & y\geqslant 0,\end{cases}$ 所以

$$P\{X_1=0,X_2=0\}=P\{Y\leqslant1,Y\leqslant2\}=P\{Y\leqslant1\}=1-e^{-1},$$

$$P\{X_1=0,X_2=1\}=P\{Y\leqslant1,Y>2\}=0,$$

$$P\{X_1=1,X_2=0\}=P\{Y>1,Y\leqslant2\}=P\{1<Y\leqslant2\}$$
$$=1-e^{-2}-(1-e^{-1})=e^{-1}-e^{-2},$$

$$P\{X_1=1,X_2=1\}=P\{Y>1,Y>2\}=P\{Y>2\}=e^{-2}.$$

(2) X_k 服从 0-1 分布，故

$$E(X_k)=e^{-k},\quad E(X_1+X_2)=e^{-1}+e^{-2}.$$

$X_2 \backslash X_1$	0	1
0	$1-e^{-1}$	$e^{-1}-e^{-2}$
1	0	e^{-2}

X_k	0	1
p_k	$1-e^{-1}$	e^{-1}

19. 两台同样的自动记录仪，每台的无故障工作时间服从参数为 5 的指数分布；首先开动其中一台，当其发生故障时停用而另一台自行启动. 试求两台记录仪无故障工作总时间 T 的概率密度 $f(x)$、数学期望和方差.　　(1997 年三)

解 以 X_i $(i=1,2)$表示两台记录仪的无故障工作时间，即

$$T=X_1+X_2.$$

X_i 的概率密度为

$$p_i(x)=\begin{cases}5e^{-5x}, & x>0,\\ 0, & x\leqslant 0.\end{cases}$$

X_1 与 X_2 相互独立，由卷积公式知

$$f(t)=\int_{-\infty}^{+\infty}p_1(x)p_2(t-x)dx=\int_0^t 25e^{-5x}\cdot e^{-5(t-x)}dx$$

$$=25e^{-5t}\int_0^t dx=25te^{-5t}.$$

所以
$$f(t)=\begin{cases}25te^{-5t}, & t>0,\\ 0, & t\leqslant 0.\end{cases}$$

$$E(X_i)=1/5,\quad D(X_i)=1/25.$$

$$E(T)=E(X_1+X_2)=2/5,\quad D(T)=D(X_1)+D(X_2)=2/25.$$

20. 游客乘电梯从底层到电视塔顶层观光，电梯于每个整点的第 5 分钟、25 分钟和 55 分钟从底层起行. 假设一游客从早八时的第 X 分钟到达底层候梯，且 X 在[0,60]上均匀分布，求该游客等候时间的数学期望. (1997 年三)

解
$$f(x)=\begin{cases}1/60, & 0\leqslant x\leqslant 60,\\ 0, & \text{其它},\end{cases}$$

而游客等待时间 Y 是 X 的函数，即

$$Y=g(X)=\begin{cases}5-X, & 0\leqslant X\leqslant 5,\\ 25-X, & 5<X\leqslant 25,\\ 55-X, & 25<X\leqslant 55,\\ 60-X+5, & 55<X\leqslant 60,\end{cases}$$

故
$$E(Y)=E[g(X)]=\int_{-\infty}^{+\infty}g(x)f(x)dx$$

$$=\int_0^5\frac{5-x}{60}dx+\int_5^{25}\frac{25-x}{60}dx$$

$$+\int_{25}^{55}\frac{55-x}{60}\mathrm{d}x+\int_{55}^{60}(60-x+5)\mathrm{d}x$$
$$=(12.5+200+400+37.5)/60=11.67.$$

21. 某流水生产线上每个产品不合格的概率为 p $(0<p<1)$，各产品合格与否相互独立，当出现一个不合格产品时即停机检修. 设开机后第一次停机时已生产了的产品个数为 X，求 X 的数学期望 $E(X)$ 和方差 $D(X)$. (2000 年一)

解 记 $q=1-p$，则 X 的概率分布为

$$P\{X=i\}=q^{i-1}p,\quad i=1,2,\cdots,$$

X 的数学期望为

$$E(X)=\sum_{i=1}^{\infty}iq^{i-1}p=p\sum_{i=1}^{\infty}iq^{i-1}=p\Big(\sum_{i=1}^{\infty}q^{i}\Big)'$$
$$=p\Big(\frac{q}{1-q}\Big)'=\frac{1}{p}.$$

又
$$E(X^2)=\sum_{i=1}^{\infty}i^2q^{i-1}p=p\sum_{i=1}^{\infty}i^2q^{i-1}=p\Big[q\Big(\sum_{i=1}^{\infty}q^{i}\Big)'\Big]'$$
$$=p\Big[\frac{q}{(1-q)^2}\Big]'=\frac{2-p}{p^2},$$

故
$$D(X)=E(X^2)-[E(X)]^2=\frac{2-p}{p^2}-\frac{1}{p^2}=\frac{1-p}{p^2}.$$

(二) 随机变量函数的数字特征

1. 设 $X_1,X_2,\cdots,X_n$ $(n>2)$ 为相互独立且同分布的随机变量，并均服从 $N(0,1)$，记 $\overline{X}=\frac{1}{n}\sum\limits_{i=1}^{n}X_i$，$Y_i=X_i-\overline{X}$，$i=1,2,\cdots,n$. 求：

(1) Y_i 的方差 $D(Y_i)$，$i=1,2,\cdots,n$；

(2) Y_1 与 Y_n 的协方差 $\mathrm{cov}(Y_1,Y_n)$；

(3) $P\{Y_1+Y_n\leqslant 0\}$. (2005 年一)

解 (1) $D(Y_i)=D(X_i-\overline{X})=D\Big[\Big(1-\frac{1}{n}\Big)X_i-\frac{1}{n}\sum\limits_{k\neq i}X_k\Big]$

$$=(n-1)/n,\quad i=1,2,\cdots,n.$$

(2) $\text{cov}(Y_1,Y_n)=E[Y_1-E(Y_1)][(Y_n-E(Y_n)]$

$$=E(X_1-\overline{X})(X_n-\overline{X})$$

$$=E(X_1X_n)+E(\overline{X}^2)-E(X_1\overline{X})-E(X_n\overline{X})$$

$$=E(X_1)E(X_n)+D(\overline{X})-\frac{1}{n}E(X_1^2)-\frac{1}{n}\sum_{k=2}^{n}E(X_1X_k)$$

$$-\frac{1}{n}E(X_n^2)-\frac{1}{n}\sum_{k=1}^{n-1}E(X_kX_n)$$

$$=-1/n.$$

(3) $Y_1+Y_n=X_1-\overline{X}+X_n-\overline{X}$

$$=\frac{n-2}{n}X_1-\frac{2}{n}\sum_{i=2}^{n-1}X_i+\frac{n-2}{n}X_n.$$

上式是相互独立的正态随机变量的线性组合，所以 Y_1+Y_n 服从正态分布.

由于 $E(Y_1+Y_n)=0$，故 $P\{Y_1+Y_n\leqslant 0\}=1/2$.

在 2005 年数学一和数学三中，此题的题设为：$X_1,X_2,\cdots,X_n$ $(n>2)$ 为来自总体 $N(0,\sigma^2)$ 的简单随机样本，$\overline{X}$ 为样本均值，因而：(1) $D(Y_i)=(n-1)/n$，$i=1,2,\cdots,n$；(2) $\text{cov}(Y_1,Y_n)=-\sigma^2/n$；数学三中的(3)为，若 $C(Y_1+Y_n)^2$ 是 σ^2 的无偏估计量，求常数 C. 于是

$$E[C(Y_1+Y_n)^2]$$

$$=CD(Y_1+Y_n)$$

$$=C[D(Y_1)+D(Y_n)+2\text{cov}(Y_1,Y_n)]$$

$$=C\left(\frac{n-1}{n}+\frac{n-1}{n}-\frac{2}{n}\right)\sigma^2=\frac{2(n-2)}{n}C\sigma^2=\sigma^2,$$

故
$$C=\frac{n}{2(n-2)}.$$

2. 设 A,B 为随机事件，且 $P(A)=1/4$，$P(B|A)=1/3$，$P(A|B)=1/2$，令

$$X=\begin{cases}1, & A\text{ 发生},\\0, & A\text{ 不发生},\end{cases}\quad Y=\begin{cases}1, & B\text{ 发生},\\0, & B\text{ 不发生},\end{cases}$$

求：(1) 二维随机变量(X,Y)的概率分布；

(2) X 与 Y 的相关系数 ρ_{XY}；

(3) $Z=X^2+Y^2$ 的概率分布.（数学一不考.）

（2004 年一、三、四）

解 (1) 由 $P(AB)=P(A)P(B|A)=1/12$,

$$P(B)=\frac{P(AB)}{P(A|B)}=\frac{1}{6}$$

得
$$P\{X=1,Y=1\}=P(AB)=1/12,$$
$$P\{X=1,Y=0\}=P(A\overline{B})=P(A)-P(AB)=1/6,$$
$$P\{X=0,Y=1\}=P(\overline{A}B)=P(B)-P(AB)=1/12,$$
$$P\{X=0,Y=0\}=P(\overline{A}\,\overline{B})=1-1/12-1/6-1/12=2/3,$$

即

X \ Y	0	1
0	2/3	1/12
1	1/6	1/12

(2) 因为

X	0	1
p_k	3/4	1/4

Y	0	1
p_k	5/6	1/6

所以
$$E(X)=1/4,\quad D(X)=3/16,$$
$$E(Y)=1/6,\quad D(Y)=5/36,\quad E(XY)=1/12,$$
$$\operatorname{cov}(X,Y)=E(XY)-E(X)E(Y)=1/24,$$

从而
$$\rho_{XY}=\frac{\operatorname{cov}(X,Y)}{\sqrt{D(X)}\sqrt{D(Y)}}=\frac{\sqrt{15}}{15}.$$

(3) Z 的可能取值为0,1,2,而

$$P\{Z=0\}=P\{X=0,Y=0\}=2/3,$$
$$P\{Z=1\}=P\{X=0,Y=1\}+P\{X=1,Y=0\}=1/4,$$
$$P\{Z=2\}=P\{X=1,Y=1\}=1/12,$$

故 Z 的分布律为

Z	0	1	2
p_k	2/3	1/4	1/12

3. 设随机变量 X 和 Y 的相关系数为 0.5，$E(X)=E(Y)=0$，$E(X^2)=E(Y^2)=2$，则 $E(X+Y)^2=$ ________. （2003 年四）

解 因为 $D(X)=E(X^2)-[E(X)]^2=2$，

$$D(Y)=2,\quad \rho_{XY}=0.5,$$

所以 $\operatorname{cov}(X,Y)=\rho_{XY}\sqrt{D(X)}\sqrt{D(Y)}=0.5\times 2=1.$

又 $\operatorname{cov}(X,Y)=E(XY)-E(X)E(Y)\Longrightarrow E(XY)=1$，

故 $E(X+Y)^2=E(X^2)+E(Y^2)+2E(XY)=2+2+2=6.$

4. 设随机变量 X 的概率密度为

$$f(x)=\begin{cases}\dfrac{1}{2}\cos\dfrac{x}{2}, & 0\leqslant x\leqslant \pi,\\ 0, & 其它,\end{cases}$$

对 X 独立地重复观察 4 次，用 Y 表示观察值大于 $\pi/3$ 的次数，求 Y^2 的数学期望. （2002 年一）

解 因为

$$P\left\{X>\frac{\pi}{3}\right\}=\int_{\pi/3}^{\pi}\frac{1}{2}\cos\frac{x}{2}\mathrm{d}x=\sin\frac{x}{2}\Big|_{\pi/3}^{\pi}=\frac{1}{2},$$

所以，$Y\sim B(4,1/2)$，于是

$$E(Y)=np=4\times 1/2=2,$$

$$D(Y)=np(1-p)=4\times 1/2\times 1/2=1,$$

得 $$E(Y^2)=[E(Y)]^2+D(Y)=2^2+1=5.$$

或由

Y	0	1	2	3	4
p_k	1/16	4/16	6/16	4/16	1/16

得 $$E(Y^2)=\frac{1}{16}\left(0\times\frac{1}{16}+1\times\frac{4}{16}+4\times\frac{6}{16}+9\times\frac{4}{16}+16\times\frac{1}{16}\right)$$
$$=5.$$

5. 设二维随机变量 (X,Y) 服从二维正态分布，则随机变量 ξ

$=X+Y$ 与 $\eta=X-Y$ 不相关的充分必要条件为(　　).

(A) $E(X)=E(Y)$;

(B) $E(X^2)-[E(X)]^2=E(Y^2)-[E(Y)]^2$;

(C) $E(X^2)=E(Y^2)$;

(D) $E(X^2)+[E(X)]^2=E(Y^2)+[E(Y)]^2$.　　(2000 年一)

解　选(B). 因为 ξ 与 η 不相关,则 $\operatorname{cov}(\xi,\eta)=0$,即

$$E\{[(X+Y)-E(X+Y)][(X-Y)-E(X-Y)]\}=0,$$

而上式左边为

$$E[(X^2-Y^2)-E(X+Y)E(X-Y)]$$
$$=E(X^2)-E(Y^2)-[E(X)]^2+[E(Y)]^2,$$

从而得　$E(X^2)-E(Y^2)-[E(X)]^2+[E(Y)]^2=0$,

即　$E(X^2)-[E(X)]^2=E(Y^2)-[E(Y)]^2$.

6. 设随机变量 X 在区间 $[-1,2]$ 上服从均匀分布,随机变量

$$Y=\begin{cases}1, & X>0,\\ 0, & X=0,\\ -1, & X<0,\end{cases}$$

则方差 $D(Y)=$________.　　(2000 年三、四)

解　由于 X 服从 $[-1,2]$ 上的均匀分布,故 Y 的分布律为

Y	1	0	−1
p_k	2/3	0	1/3

$$D(Y)=E(Y^2)-[E(Y)]^2=1-(1/3)^2=8/9.$$

7. 设 X 表示 10 次独立重复射击命中目标的次数,每次射中目标的概率为 0.4,则 X^2 的数学期望 $E(X^2)=$________.

(1995 年一)

解　因为　$X\sim B(10,0.4)$,

$$E(X)=np=4,\quad D(X)=np(1-p)=2.4,$$

所以　$E(X^2)=D(X)+[E(X)]^2=2.4+4^2=18.4$.

8. 设随机变量 X 服从参数为 λ 的泊松分布,且已知

$E[(X-1)(X-2)]=1$，则 $\lambda=$ ________.　　(1999 年四)

解　$E(X)=\lambda, D(X)=\lambda$，而

$$\begin{aligned} E[(X-1)(X-2)] &= E(X^2)-3E(X)+2 \\ &= D(X)+[E(X)]^2-3E(X)+2 \\ &= \lambda+\lambda^2-3\lambda+2=1, \end{aligned}$$

解方程，得 $\lambda=1$.

9. 设一次试验成功的概率为 p，进行 100 次独立重复试验，当 $p=$ ________ 时，成功次数的标准差为最大，最大值 σ 为 ________.

(1998 年四)

解　成功次数 $X\sim B(100,p)$，$D(X)=np-np^2$. 由

$$\mathrm{d}D(X)/\mathrm{d}p=n-2np,$$

令上式为零得 $p=1/2$. 此时

$$D(X)=100\times 1/2\times(1-1/2)=25,\quad \sigma=5.$$

10. 假设一部机器在一天内发生故障的概率为 0.2，机器发生故障时全天停止工作，若一周五个工作日无故障，可获利润 10 万元；发生一次故障仍可获利润 5 万元；发生两次故障获利润为 0；发生三次或三次以上故障亏本 2 万元. 求一周内利润期望值是多少？

(1996 年五)

解　以 X 表示一周内无故障工作天数，则 $X\sim B(5,0.2)$，$P\{X=k\}=C_5^k\times 0.2^k\times 0.8^{5-k}$，$k=0,1,\cdots,5$. X 的分布律为

X	0	1	2	$\geqslant 3$
p_k	0.328	0.410	0.205	0.057
Y	10	5	0	-2

故

$$\begin{aligned} E(Y) &= 0.328\times 10+0.410\times 5+0.057\times(-2) \\ &= 5.216\ (\text{万元}). \end{aligned}$$

11. 一商店经销某种商品，每周进货的数量 X 与顾客对该种商品的需求量 Y 是相互独立的随机变量，且都服从区间[10,20]上的均匀分布，商店每售出一单位商品可得利润 1000 元. 若需求量

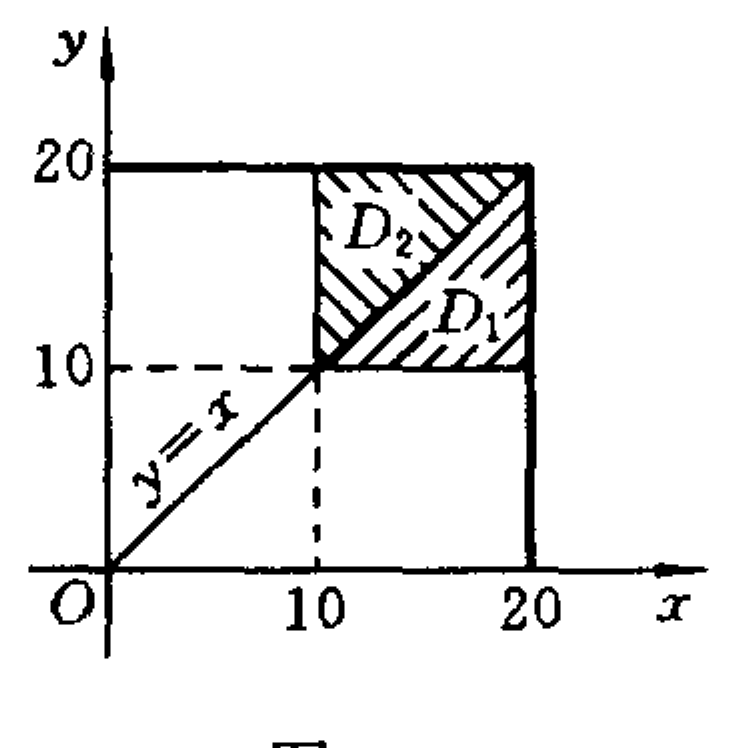


图 4.6

超过了进货量，商店可以从其它商店调剂供应，这时每单位商品获利润500元. 试计算此商店经销该种商品每周所得利润的期望值.

(1998 年三)

解 如图 4.6 所示，设商店每周所得利润为Q，则

$$Q=\begin{cases}1000Y, & Y\leqslant X,\\ 1000X+500(Y-X), & Y>X,\end{cases}$$

由于X,Y均服从均匀分布$U(10,20)$，故联合密度为

$$f(x,y)=\begin{cases}1/10\times 1/10=1/100, & 10\leqslant x,y\leqslant 20,\\ 0, & 其它.\end{cases}$$

从而，利润的期望值

$$\begin{aligned}E(Q)&=\iint\limits_{D_1}1000y\times\frac{1}{100}\mathrm{d}x\mathrm{d}y+\iint\limits_{D_2}500(x+y)\times\frac{1}{100}\mathrm{d}x\mathrm{d}y\\ &=10\int_{10}^{20}\mathrm{d}y\int_{y}^{20}y\mathrm{d}x+5\int_{10}^{20}\mathrm{d}y\int_{10}^{y}(x+y)\mathrm{d}x\\ &=14166.67\ (元).\end{aligned}$$

12. 设某种商品的每周需求量X是服从区间[10,30]上的均匀分布的随机变量，而经销商店进货数量为[10,30]中的某一整数. 商店每销售一单位商品可获利500元；若供大于求则削价处理，每处理一单位商品亏损100元；若供不应求则可从外部调剂供应，此时每一单位商品仅获利300元. 为使商品所获利润期望值不少于9280元，试确定最少进货量. (1998 年四)

解 设进货量为a，则利润为

$$M_a=\begin{cases}500a+(X-a)300\\ 500X-(a-X)100\end{cases}=\begin{cases}300X+200a, & a<X\leqslant 30,\\ 600X-100a, & 10\leqslant X\leqslant a,\end{cases}$$

利润期望为

$$E(M_a)=\int_{10}^{30}M_a\times\frac{1}{20}\mathrm{d}x$$

$$=\frac{1}{20}\int_{10}^{a}(600X-100a)\mathrm{d}x+\frac{1}{20}\int_{a}^{30}(300x+200a)\mathrm{d}x$$
$$=-7.5a^2+350a+5250.$$

由于$E(M_a)\geqslant 9280$,则由

$$-7.5a^2+350a+5250\geqslant 9280$$

可解得$62/3\leqslant a\leqslant 26$,故知利润期望值不少于 9280 元的最少进货量为 21 单位.

13. 假设由自动线加工的某种零件的内径(单位:mm)X 服从正态分布$N(\mu,1)$,内径小于10 或大于12 的为不合格品,其余为合格品.销售每件合格品获利,销售每件不合格品亏损,已知销售利润(单位:元)T 与销售零件的内径X 有如下关系:

$$T=\begin{cases}-1, & X<10,\\ 20, & 10\leqslant X\leqslant 12,\\ -5, & X>12.\end{cases}$$

问:平均内径μ 取何值时,销售一个零件的平均利润最大?

(1994 年四)

解 由题设条件知,平均利润为

$$\begin{aligned}E(T)&=20P\{10\leqslant X\leqslant 12\}+(-1)P\{X<10\}\\&\quad+(-5)P\{X>12\}\\&=20[\Phi(12-\mu)-\Phi(10-\mu)]-\Phi(10-\mu)\\&\quad-5[1-\Phi(12-\mu)]\\&=25\Phi(12-\mu)-21\Phi(10-\mu)-5,\end{aligned}$$

其中$\Phi(x)$是标准正态分布函数.设$\varphi(x)$为标准正态密度,则有

$$\frac{\mathrm{d}E(T)}{\mathrm{d}\mu}=-25\varphi(12-\mu)+21\varphi(10-\mu),$$

令上式等于零,得

$$-25/\sqrt{2\pi}\mathrm{e}^{-(12-\mu)^2/2}+21/\sqrt{2\pi}\mathrm{e}^{-(10-\mu)^2/2}=0,$$

即 $$25\mathrm{e}^{-(12-\mu)^2}=21\mathrm{e}^{-(10-\mu)^2/2}.$$

得 $$\mu=\mu_0=11-1/2\times\ln(25/21)=10.9.$$

即，当$\mu=\mu_0=10.9$ mm 时，平均利润最大.

14. 假设二维随机变量(X,Y)在矩形$G=\{(x,y)\mid 0\leqslant x\leqslant 2,0\leqslant y\leqslant 1\}$上服从均匀分布，记

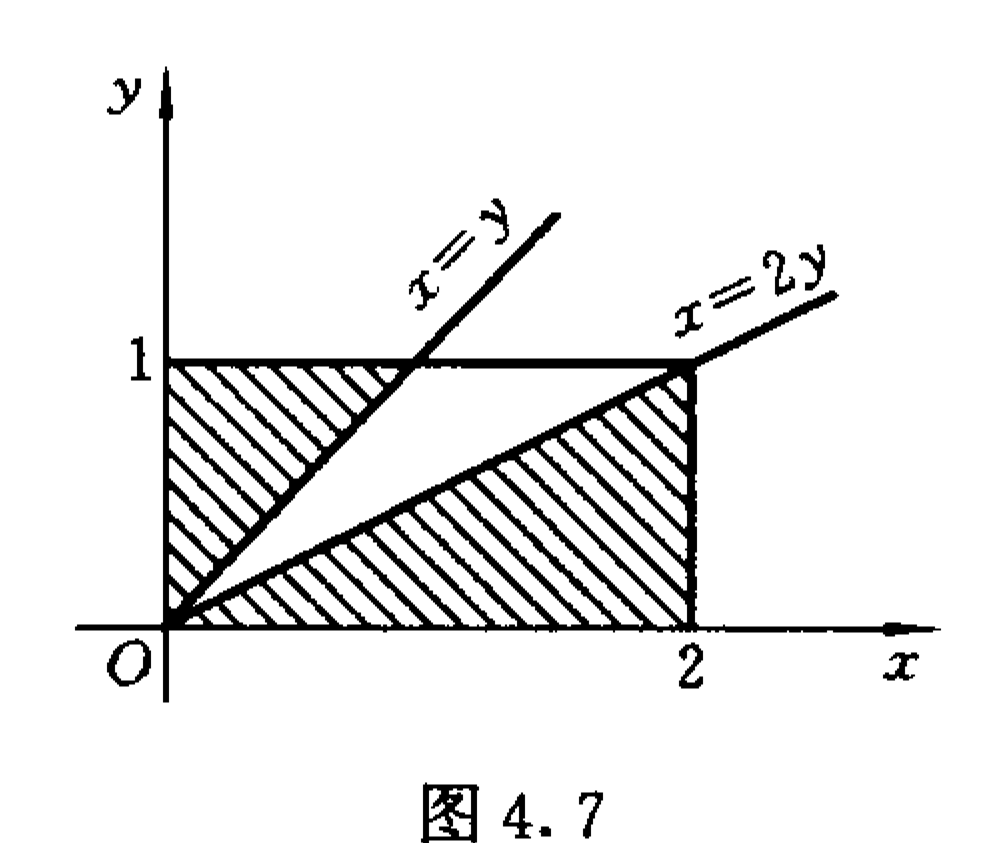


图 4.7

$$U=\begin{cases}0, & X\leqslant Y,\\ 1, & X>Y,\end{cases}$$

$$V=\begin{cases}0, & X\leqslant 2Y,\\ 1, & X>2Y.\end{cases}$$

(1) 求U与V的联合分布；

(2) 求U和V的相关系数.

(1999 年三)

解 由图 4.7 可知

$$P\{X\leqslant Y\}=1/4,\quad P\{X>2Y\}=1/2,$$
$$P\{Y<X\leqslant 2Y\}=1/4.$$

(1) (U,V)有四种取值：$(0,0),(0,1),(1,0),(1,1)$.

$$P\{U=0,V=0\}=P\{X<Y,X\leqslant 2Y\}=P\{X\leqslant Y\}=1/4,$$
$$P\{U=0,V=1\}=P\{X\leqslant Y,X>2Y\}=0,$$
$$P\{U=1,V=0\}=P\{X>Y,X\leqslant 2Y\}=P\{Y<X\geqslant 2Y\}=1/4,$$
$$P\{U=1,V=1\}=P\{X>Y,X>2Y\}=P\{X>2Y\}=1/2.$$

(2) 由题(1)与图 4.7 可知

$$UV\sim\begin{pmatrix}0 & 1\\ 1/2 & 1/2\end{pmatrix},\quad U\sim\begin{pmatrix}0 & 1\\ 1/4 & 3/4\end{pmatrix},\quad V\sim\begin{pmatrix}0 & 1\\ 1/2 & 1/2\end{pmatrix},$$

故
$$E(U)=3/4,\quad D(U)=3/16,\quad E(V)=1/2,$$
$$D(V)=1/4,\quad E(UV)=1/2,$$
$$\operatorname{cov}(U,V)=E(UV)-E(U)E(V)=1/8,$$
$$\rho_{UV}=\operatorname{cov}(U,V)/[\sqrt{D(U)}\sqrt{D(V)}]=1/\sqrt{3}.$$

15. 设随机变量X的概率分布密度为

$$f(x)=\frac{1}{2}e^{-|x|},\quad -\infty<x<+\infty.$$

(1) 求X的数学期望$E(X)$和方差$D(X)$；

(2) 求 X 与 $|X|$ 的协方差，并问：X 与 $|X|$ 是否不相关？

(3) 问：X 与 $|X|$ 是否相互独立？为什么？ (1993 年一)

解 (1) 由定义，有

$$E(X)=\int_{-\infty}^{+\infty}\frac{x\mathrm{e}^{-|x|}}{2}\mathrm{d}x=0,\quad D(X)=\int_{0}^{+\infty}x^2\mathrm{e}^{-x}\mathrm{d}x=2.$$

(2) $$\begin{aligned}\operatorname{cov}(X,|X|)&=E\{[X-E(X)][|X|-E(Y)]\}\\&=E(X|X|)-E(X)E(|X|)\\&=E(X|X|)=\int_{-\infty}^{+\infty}x|x|f(x)\mathrm{d}x=0,\end{aligned}$$

所以，X 与 $|X|$ 不相关.

(3) 对给定的 $0<a<+\infty$，显然事件 $\{|X|<a\}$ 包含在事件 $\{X<a\}$ 内，且 $P\{X<a\}<1, 0<P\{|X|<a\}$，故

$$P\{X<a,|X|<a\}=P\{|X|<a\}.$$

但 $$P\{X<a\}P\{|X|<a\}<P\{|X|<a\},$$

所以 $$P\{X<a,|X|<a\}\neq P\{X<a\}P\{|X|<a\},$$

因此，X 与 $|X|$ 不相互独立.

16. 已知随机变量 (X,Y) 服从二维正态分布，且 X 和 Y 分别服从正态分布 $N(1,3^2)$ 和 $N(0,4^2)$，X 和 Y 的相关系数 $\rho_{XY}=-1/2$. 设 $Z=X/3+Y/2$，

(1) 求 Z 的数学期望 $E(Z)$ 和方差 $D(Z)$；

(2) 求 X 与 Z 的相关系数 ρ_{XZ}；

(3) 问：X 与 Z 是否相互独立？为什么？ (1994 年一)

解 (1) $E(Z)=E(X)/3+E(Y)/2=1/3$，

$$\begin{aligned}D(Z)&=D(X)/9+D(Y)/4+2\rho_{XY}\sqrt{D(X)}/3\cdot\sqrt{D(Y)}/2\\&=1+4-2=3.\end{aligned}$$

(2) $$\begin{aligned}\operatorname{cov}(X,Z)&=\operatorname{cov}(X,X)/3+\operatorname{cov}(X,Y)\\&=3^2/3+(-1/2)\times3\times4/2=0,\end{aligned}$$

所以 $$\rho_{XZ}=\operatorname{cov}(X,Z)/[\sqrt{D(X)}\sqrt{D(Z)}]=0.$$

(3) 因为 X,Y 均为正态分布，故 Z 也为正态分布. 由 $\rho_{XZ}=0$，

所以 X 与 Z 相互独立.

17. 某箱装有100件产品，其中一、二、三等品分别为80，10，10件，现在从中随机抽取一件，记

$$X_i=\begin{cases}1, & \text{抽到 } i \text{ 等品},\\ 0, & \text{其它},\end{cases}\quad i=1,2,3.$$

试求：(1) 随机变量 X_1 与 X_2 的联合分布；

(2) 随机变量 X_1 与 X_2 的相关系数. （1998年四）

解 (1) 以 A_i $(i=1,2,3)$ 表示抽到 i 等品，于是

$$P(A_i)=0.8,\quad P(A_2)=0.1,\quad P(A_3)=0.1.$$

于是
$$P\{X_1=0,X_2=0\}=P\{A_3\}=0.1,$$
$$P\{X_1=0,X_2=1\}=P(A_2)=0.1,$$
$$P\{X_1=1,X_2=0\}=P(A_1)=0.8,$$
$$P\{X_1=1,X_2=1\}=P(\varnothing)=0.$$

联合分布律为

X_2 \ X_1	0	1
0	0.1	0.8
1	0.1	0

(2)
$$E(X_1)=0.8,\quad E(X_2)=0.1,$$
$$E(X_1^2)=0.8,\quad E(X_2^2)=0.1,$$
$$D(X_1)=0.8-0.8^2=0.16,\quad D(X_2)=0.1-0.1^2=0.09,$$
$$E(X_1X_2)=0,$$
$$\operatorname{cov}(X_1,X_2)=E(X_1,X_2)-E(X_1)E(X_2)=0-0.08=-0.08,$$
$$\rho_{X_1X_2}=\operatorname{cov}(X_1,X_2)/[\sqrt{D(X_1)}\sqrt{D(X_2)}]$$
$$=-0.08/(\sqrt{0.16}\sqrt{0.09})=-2/3.$$

18. 设 A,B 是二随机事件，随机变量

$$X=\begin{cases}1, & \text{若 } A \text{ 出现},\\ -1, & \text{若 } A \text{ 不出现};\end{cases}\qquad Y=\begin{cases}1, & \text{若 } B \text{ 出现},\\ -1, & \text{若 } B \text{ 不出现},\end{cases}$$

试证明随机变量 X 和 Y 不相关的充分必要条件是 A 与 B 相互独立.　　　　　　　　　　　　　　　　　　　　　　（2000 年三、四）

证　记 $P(A)=p_1, P(B)=p_2, P(AB)=p_{12}$，则

$$E(X)=P(A)-P(\overline{A})=p_1-(1-p_1)=2p_1-1.$$

由于 XY 只有两个可能值 1 和 -1，所以

$$P\{XY=1\}=P(AB)+P(\overline{A}\overline{B})=2p_{12}-p_1-p_2+1,$$

$$P\{XY=-1\}=1-P\{XY=1\}=p_1+p_2-2p_{12},$$

故

$$E(XY)=P\{XY=1\}-P\{XY=-1\}=4p_{12}+1-2p_1-2p_2,$$

于是

$$\begin{aligned}\operatorname{cov}(X,Y)&=E(XY)-E(X)E(Y)\\&=4p_{12}+1-2p_1-2p_2-(2p_1-1)(2p_2-1)\\&=4p_{12}-4p_1p_2.\end{aligned}$$

因此，$\operatorname{cov}(X,Y)$ 当且仅当 A,B 相互独立时为零. 此时 $p_{12}=p_1p_2$，即 X 和 Y 不相关的充分必要条件是 A 与 B 相互独立.

19. 设二维随机变量 (X,Y) 的密度函数为

$$f(x,y)=\frac{1}{2}[\varphi_1(x,y)+\varphi_2(x,y)],$$

其中 $\varphi_1(x,y)$ 和 $\varphi_2(x,y)$ 都是二维正态密度函数，且它们对应的二维随机变量的相关系数为 1/3 和 $-1/3$. 它们的边缘密度函数所对应的随机变量的数学期望都是零，方差都是 1.

(1) 求随机变量 X 和 Y 的密度函数 $f_1(x), f_2(y)$ 及 X 和 Y 的相关系数 ρ_{XY}（可直接利用二维正态密度的性质）；

(2) 问：X 和 Y 是否独立？为什么？　　　　　　　　　（2000 年四）

解　(1) 二维正态分布的两个边缘分布都是一维正态分布，因此，$\varphi_1(x,y)$ 和 $\varphi_2(x,y)$ 的两个边缘密度为标准正态密度函数，故

$$f_1(x)=\int_{-\infty}^{+\infty}f(x,y)\mathrm{d}y=\frac{1}{2}\left[\int_{-\infty}^{+\infty}\varphi_1(x,y)\mathrm{d}y+\int_{-\infty}^{+\infty}\varphi_2(x,y)\mathrm{d}y\right]$$

$$=\frac{1}{2}\left(\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-x^2/2}+\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-x^2/2}\right)=\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-x^2/2},$$

同理 $$f_2(y)=\frac{1}{\sqrt{2\pi}}e^{-y^2/2}.$$

由 $X\sim N(0,1)$，$Y\sim N(0,1)$知

$$E(X)=E(Y)=0,\quad D(X)=D(Y)=1.$$

随机变量 X 和 Y 的相关系数

$$\rho_{XY}=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}xyf(x,y)\mathrm{d}x\mathrm{d}y$$
$$=\frac{1}{2}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}xy\varphi_1(x,y)\mathrm{d}x\mathrm{d}y+\frac{1}{2}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}xy\varphi_2(x,y)\mathrm{d}x\mathrm{d}y$$
$$=\frac{1}{2}\left(\frac{1}{3}-\frac{1}{3}\right)=0.$$

(2) 由题给条件得

$$f(x,y)=\frac{3}{8\pi\sqrt{2}}\left[e^{-9(x^2-2xy/3+y^2)/16}+e^{-9(x^2+2xy/3+y^2)/16}\right]$$

$$f_1(x)f_2(y)=\frac{1}{2\pi}e^{-x^2/2}\cdot e^{-y^2/2}=\frac{1}{2\pi}e^{-(x^2+y^2/2)}$$

知 $$f(x,y)\neq f_1(x)f_2(y),$$

所以，X 与 Y 不相互独立.

20. 设二维随机变量(X,Y)在区域$D:0<x<1,|y|<x$ 内服从均匀分布，求关于X 的边缘概率密度及随机变量 $Z=2X+1$ 的方差$D(Z)$.　　(1990 年一)

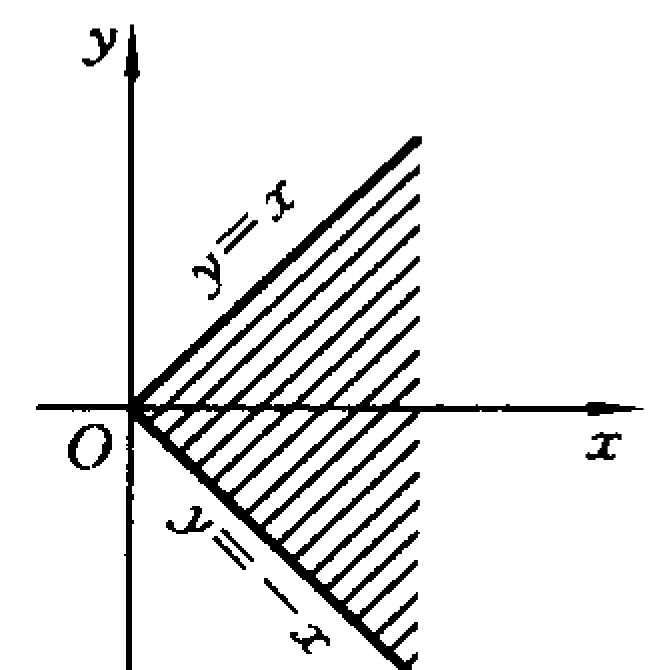


图 4.8

解　如图4.8所示，因为

$$f(x,y)=\begin{cases}1, & (x,y)\in G,\\ 0, & 其它,\end{cases}$$

所以 $$f_X(x)=\begin{cases}\int_{-x}^{x}\mathrm{d}y=2x, & x<x<1,\\ 0, & 其它,\end{cases}$$

$$D(X)=D(2X+1)=4D(X)=4\int_0^1 x^2\times 2x\mathrm{d}x-4\left(\int_0^1 x\mathrm{d}x\right)^2$$
$$=2-4\times 4/9=2/9.$$

21. 将一枚硬币重复掷n 次，以X 和Y 表示正面向上和反面向上的次数，则 X 和 Y 的相关系数等于(　　).

(A) -1；　　(B) 0；　　(C) 1/2；　　(D) 1.

(1999 年一)

解　选(A)，因为 $Y=n-X$，是线性负相关.

22. 设随机变量 X 和 Y 相互独立且同分布，记 $U=X-Y$，$V=X+Y$，则随机变量 U 与 V 必然(　　).

(A) 不相互独立；　　(B) 相互独立；

(C) 相关系数不为零；　　(D) 相关系数为零.

(1995 年四)

解　选(D). 因为 X，Y 相互独立且同分布，所以

$$\begin{aligned}\mathrm{cov}(U,V)&=E(UV)-E(U)E(V)\\&=E(X^2-Y^2)-E(X-Y)E(X+Y)\\&=E(X^2)-E(Y^2)-[E(X)]^2+[E(Y)]^2\\&=D(X)-D(Y)=0,\end{aligned}$$

即
$$\rho_{XY}=0.$$

23. 设随机变量 X 和 Y 的联合分布在以点(0,1)，(1,0)，(1,1)为顶点的三角形区域上服从均匀分布，试求随机变量 $U=X+Y$ 的方差.　　(2001 年四)

解　如图 4.9 所示. 因为 $S_G=2$，所以

$$f(x,y)=\begin{cases}2, & (x,y)\in G,\\ 0, & (x,y)\overline{\in} G,\end{cases}$$

$$f_X(x)=\begin{cases}\displaystyle\int_{1-x}^{1}2\mathrm{d}y=2x, & 0<x<1,\\ 0, & \text{其它}.\end{cases}$$

图 4.9

因此　$E(X)=\displaystyle\int_0^1 2x\cdot x\mathrm{d}x=2/3$，

$$E(X^2)=\int_0^1 2x\cdot x^2\mathrm{d}x=1/2,\quad D(X)=1/2-4/9=1/18.$$

同理
$$E(Y)=2/3,\quad D(Y)=1/18,$$

$$E(XY)=2\int_0^1 x\mathrm{d}x\int_{1-x}^{1}y\mathrm{d}y=5/12,$$

$$\mathrm{cov}(X,Y)=E(XY)-E(X)E(Y)=5/12-4/9=-1/36.$$

从而 $D(U)=D(X+Y)=D(X)+D(Y)+2\text{cov}(X,Y)$

$$=1/18+1/18-2/36=1/18.$$

24. 设总体 X 服从正态分布 $N(\mu,\sigma^2)$ $(\sigma>0)$，从该总体中抽取简单随机样本 $X_1,X_2,\cdots,X_{2n}$ $(n\geqslant 2)$，其样本均值为

$$\overline{X}=\frac{1}{2n}\sum_{\lambda=1}^{2n}X_i,$$

求统计量 $Y=\sum_{i=1}^{n}(X_i+X_{n+i}-2\overline{X})^2$ 的数学期望 $E(Y)$.

(2001 年一)

解 考虑 $(X_1+X_{n+1}),(X_2+X_{n+2}),\cdots,(X_n+X_{2n})$，将其视为取自总体 $N(2\mu,2\sigma^2)$ 的简单随机样本，则其样本均值为

$$\frac{1}{n}\sum_{i=1}^{n}(X_i+X_{n+i})=\frac{1}{n}\sum_{i=1}^{2n}X_i=2\overline{X},$$

样本方差为 $\frac{1}{n-1}Y$.

由于 $E\left(\frac{1}{n-1}Y\right)=2\sigma^2$，所以

$$E(Y)=(n-1)(2\sigma^2)=2(n-1)\sigma^2.$$

25. 设随机变量 X 和 Y 的联合分布律为

X \ Y	−1	0	1
0	0.07	0.18	0.15
1	0.08	0.32	0.20

则 (1) X 和 Y 的相关系数 $\rho_{XY}=$________; (2002 年四)

(2) X^2 和 Y^2 的协方差 $\text{cov}(X^2,Y^2)=$________. (2002 年三)

解 因为 X,Y 的边缘分布律为

X	0	1
p_k	0.4	0.6

Y	−1	0	1
p_k	0.15	0.5	0.35

(1) $E(X)=0.6$, $E(Y)=0.2$, $E(XY)=0.12$,

由于 $E(XY)=E(X)E(Y)$，所以 $\rho_{XY}=0$.

(2) $E(X^2)=0.6$, $E(Y^2)=0.5$, $E(X^2Y^2)=0.28$,

由于 $E(X^2Y^2)-E(X^2)E(Y^2)=0.28-0.6\times0.5=-0.02$,

所以 $$\operatorname{cov}(X^2,Y^2)=-0.02.$$

26. 假设随机变量U在区间$[-2,2]$上服从均匀分布,随机变量

$$X=\begin{cases}-1, & 若U\leqslant -1,\\ 1, & 若U>-1,\end{cases}\quad Y=\begin{cases}-1, & 若U\leqslant 1,\\ 1, & 若U>1,\end{cases}$$

试求:(1) X和Y的联合概率分布;(2) $D(X+Y)$. (2002年三)

解 (1) 随机向量(X,Y)有四个可能值:$(-1,-1)$,$(-1,1)$,$(1,-1)$,$(1,1)$.

$$P\{X=-1,Y=-1\}=P\{U\leqslant -1,U\leqslant 1\}=1/4,$$
$$P\{X=-1,Y=1\}=P\{U\leqslant -1,U>1\}=0,$$
$$P\{X=1,Y=-1\}=P\{U>-1,U\leqslant 1\}=1/2,$$
$$P\{X=1,Y=1\}=P\{U>-1,U>1\}=1/4.$$

于是,得X和Y的联合概率分布为

$$(X,Y)\sim\begin{pmatrix}(-1,-1) & (-1,1) & (1,-1) & (1,1)\\ 1/4 & 0 & 1/2 & 1/4\end{pmatrix}$$

(2) $X+Y$和$(X+Y)^2$的概率分布相应为

$$X+Y\sim\begin{pmatrix}-2 & 0 & 2\\ 1/4 & 1/2 & 1/4\end{pmatrix},\quad (X+Y)^2\sim\begin{pmatrix}0 & 4\\ 1/2 & 1/2\end{pmatrix}.$$

由此可见

$$E(X+Y)=-1/2+1/2=0,\quad E[(X+Y)^2]=2,$$
$$D(X+Y)=E[(X+Y)^2]=2.$$

第五章　大数定律与中心极限定理

第一节　大数定律

主要内容

1. 契比雪夫不等式

设随机变量 X 的数学期望 $E(X)=\mu$，方差 $D(X)=\sigma^2$，则对任意 $\varepsilon>0$，有不等式

$$P\{|X-\mu|\geqslant\varepsilon\}\leqslant\sigma^2/\varepsilon^2,$$

或

$$P\{|X-\mu|<\varepsilon\}\geqslant 1-\sigma^2/\varepsilon^2.$$

2. 大数定律

(1) 契比雪夫定理的特殊情形　设 $X_1,X_2,\cdots,X_n,\cdots$ 是相互独立的随机变量序列，有相同的数学期望和方差，$E(X_i)=\mu$，$D(X_i)=\sigma^2\quad(i=1,2,\cdots)$，则对任意给定的 $\varepsilon>0$，有

$$\lim_{n\to\infty}P\left\{\left|\frac{1}{n}\sum_{i=1}^{n}X_i-\mu\right|<\varepsilon\right\}=1.$$

(2) 契比雪夫定理　设 $X_1,X_2,\cdots,X_n,\cdots$ 是相互独立的随机变量序列，$E(X_i)$ 和 $D(X_i)$ 都存在，且 $D(X_i)\leqslant C\ (i=1,2,\cdots)$，则对任意的 $\varepsilon>0$，有

$$\lim_{n\to\infty}P\left\{\left|\frac{1}{n}\sum_{i=1}^{n}X_i-\frac{1}{n}\sum_{i=1}^{n}E(X_i)\right|<\varepsilon\right\}=1.$$

(3) 伯努利定理　设 n_A 是 n 次重复独立试验中事件 A 发生的次数，p 是事件 A 在一次试验中发生的概率，则对任意的 $\varepsilon>0$，有

$$\lim_{n\to\infty}P\left\{\left|\frac{n_A}{n}-p\right|<\varepsilon\right\}=1.$$

（4）辛钦定理　设 $X_1,X_2,\cdots,X_n,\cdots$ 是相互独立且同分布的随机变量序列，且 $E(X_i)=\mu$，则对任意 $\varepsilon>0$，有

$$\lim_{n\to\infty}P\left\{\left|\frac{1}{n}\sum_{n=1}^{\infty}X_i-\mu\right|<\varepsilon\right\}=1.$$

疑 难 解 析

1. 契比雪夫不等式有什么作用？它的意义是什么？

答　契比雪夫不等式 $P\{|X-\mu|\leqslant\varepsilon\}\geqslant1-\sigma^2/\varepsilon^2$ 反映了随机变量 X 的取值落在其数学期望 $E(X)=\mu$ 的 ε 邻域内的概率不小于 $1-\sigma^2/\varepsilon^2$. 它的意义在于：当知道随机变量 X 的数学期望与方差时，我们可以估计 X 落在以 $E(X)$ 为中心的某一区间内的概率（至少给出一个下限）.

它的作用有四个方面：(1) 估计概率. 当 $E(X)$，$D(X)$ 和 ε 给定时，可依公式直接计算概率. (2) 当概率确定时，估计所需区间的长度，即已知 $E(X)$，$D(X)$ 和 p，确定 ε 值. (3) 估计试验次数. 在 n 重伯努利试验中，频率 n_A/n 与试验次数有关，在已知 $E(X)$，$D(X)$ 和 p 时，可以用契比雪夫不等式确定 n. (4)是推导其它定理的依据.

注意，契比雪夫不等式给出的估计是十分粗糙的，精确的估计要通过其它更优秀的方法来给出.

2. 大数定律的意义是什么？

答　大数定律深刻地揭示了随机事件的概率与频率之间的关系，因此是概率论的重要理论基础. 大数定律从大量测量值的平均值出发，讨论并反映了算术平均值及频率的稳定性.

教材讲述的大数定律都是弱大数定律，它们的条件各不相同，但结论是一致的：从理论上肯定了用算术平均值代替均值，以频率代替概率的合理性. 大数定理既验证了概率论中一些假设的合理

性，又为数理统计中用样本推断总体提供了理论依据.

契比雪夫定理：$\frac{1}{n}\sum_{i=1}^{n}X_i \xrightarrow{p} \frac{1}{n}\sum_{i=1}^{n}E(X_i)$；

伯努利定理：$\frac{n_A}{n} \xrightarrow{p} p$；

辛钦定理：$\frac{1}{n}\sum_{i=1}^{n}X_i \xrightarrow{p} \mu$.

3. 依概率收敛的意义是什么？

答 依概率收敛即依概率1收敛. 其定义是：设随机变量序列 $X_1, X_2, \cdots, X_n$，对任意 $\varepsilon>0$，有

$$\lim_{n\to\infty}P\{|Y_n-a|<\varepsilon\}=1 \ (a\ 为常数),$$

则称序列 $Y_1, Y_2, \cdots, Y_n, \cdots$ 依概率收敛于 a，记为

$$Y_n \xrightarrow{p} a.$$

依概率收敛与微积分中的收敛的不同在于：微积分中的收敛是确定的，即对任给的 $\varepsilon>0$，当 $n>N$ 时，必有 $|x_n-a|<\varepsilon$ 成立. 而依概率收敛是，对任给的 $\varepsilon>0$，当 n 很大时，事件 $\{|x_n-a|<\varepsilon\}$ 发生的概率为1，但不排除偶然事件 $\{|x_n-a|\geqslant\varepsilon\}$ 的发生.

方法、技巧与典型例题分析

一、契比雪夫不等式及应用

利用契比雪夫不等式解题，常见的题型是估计随机变量 X 在某区间内的概率，估计区间长度，估计试验次数. 解题的步骤是：首先确定恰当的随机变量 X，计算 $E(X)$ 与 $D(X)$；其次确定 $\varepsilon>0$ 的值；最后，由契比雪夫不等式（两种形式可根据需要选择一种）进行计算，其技巧主要表现在计算 $E(X)$，$D(X)$ 与分解随机变量 X 上.

例1 若随机变量 X 服从参数为2的泊松分布，用契比雪夫不等式估计，$P\{|X-2|\geqslant4\}=$ ________.

解 若 $X\sim\pi(2)$，$E(X)=D(X)=2$，由契比雪夫不等式，有

$$P\{|X-2|\geqslant 4\}=P\{|X-\mu|\geqslant 4\}\geqslant 2/4^2=1/8.$$

例2 设 $X_1,X_2,\cdots,X_n$ 是相互独立且同分布的随机变量，$E(X_i)=\mu,D(X_i)=8\ (i=1,2,\cdots,n)$，求 $\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$ 所满足的契比雪夫不等式，并估计 $P\{|\overline{X}-\mu|<4\}\geqslant\alpha$ 中的 α.

解 先求 $E(\overline{X})$ 和 $D(\overline{X})$. 因为

$$E(\overline{X})=E\left\{\frac{1}{n}\sum_{i=1}^{n}X_i\right\}=\frac{1}{n}\sum_{i=1}^{n}E(X_i)=\frac{1}{n}n\mu=\mu,$$

$$D(\overline{X})=D\left\{\frac{1}{n}\sum_{i=1}^{n}X_i\right\}=\frac{1}{n^2}\sum_{i=1}^{n}D(X_i)=\frac{1}{n^2}nD(X_i)=\frac{8}{n},$$

所以，满足 $\overline{X}$ 的契比雪夫不等式为

$$P\{|\overline{X}-\mu|\geqslant\varepsilon\}\leqslant\frac{D(\overline{X})}{\varepsilon^2}=\frac{8}{n\varepsilon^2}.$$

当 $\varepsilon=4$ 时，即为

$$P\{|\overline{X}-\mu|<4\}\geqslant 1-\frac{8}{4^2n}=1-\frac{1}{2n}.$$

例3 若随机变量 X 服从 $[-1,b]$ 上的均匀分布，且由契比雪夫不等式得 $P\{|X-1|<\varepsilon\}\geqslant 2/3$，则 $b=$ ________，$\varepsilon=$ ________.

解 $E(X)=(-1+b)/2,\quad D(X)=(b+1)^2/12,$

又 $$P\{|X-E(X)|<\varepsilon\}\geqslant D(X)/\varepsilon^2,$$

比照 $$P\{|X-1|<\varepsilon\}\geqslant 2/3,$$

得 $$E(X)=1=(-1+b)/2\Longrightarrow b=3,$$

故 $D(X)=16/12=4/3\Longrightarrow 1-D(X)/\varepsilon^2=2/3\Longrightarrow\varepsilon^2=2.$

例4 随机地掷6枚骰子，利用契比雪夫不等式估计6枚骰子出现点数之和在15点到27点之间的概率.

解 以 $X_i\ (i=1,2,\cdots,6)$ 记第 i 枚骰子出现的点数，显然 X_i 相互独立，6枚骰子出现点数的总和 $X=\sum_{i=1}^{n}X_i$，所以

$$E(X_i)=\frac{1}{6}(1+2+3+4+5+6)=\frac{21}{6},$$

$$D(X_i)=\frac{1}{6}\left[\left(1-\frac{21}{6}\right)^2+\left(2-\frac{21}{6}\right)^2+\cdots+\left(6-\frac{21}{6}\right)^2\right]=\frac{35}{12}.$$

故 $$E(X)=21,\quad D(X)=35/2.$$

由契比雪夫不等式，有

$$P\{15<X<27\}=P\{|X-2|<6\}\geqslant 1-\frac{35}{2}\Big/6^2=\frac{37}{72}.$$

例5 设随机变量 X 的分布律为

X	1	2	3
p_k	0.3	0.5	0.2

试用契比雪夫不等式估计概率 $P\{|X-E(X)|\geqslant 1\}$.

解 $E(X)=1\times 0.3+2\times 0.5+3\times 0.2=1.9$,

$$E(X^2)=1\times 0.3+4\times 0.5+9\times 0.2=4.1,$$

$$D(X)=E(X^2)-[E(X)]^2=0.49,$$

所以 $$P\{|X-E(X)|\geqslant 1\}\leqslant 0.49/1=0.49.$$

而按概率计算，则有

$$\begin{aligned}P\{|X-1.9|\geqslant 1\}&=1-P\{|X-1.9|<1\}\\&=1-P\{0.9<X<2.9\}\\&=1-P\{1\}-P\{2\}=0.2.\end{aligned}$$

可见契比雪夫不等式的估计是很粗糙的.

例6 设在每次试验中，事件 A 发生的概率 $p=1/4$.

(1) 进行 300 次重复独立试验，以 X 记 A 发生的次数，用契比雪夫不等式估计 X 与 $E(X)$ 的偏差不大于 50 的概率；

(2) 问：是否可用 0.925 的概率确信，在 1000 次试验中，A 发生的次数在 200 到 300 之间？

解 这是一个问题的正反两种不同的提法.

(1) 由 $X\sim B(300,1/4)$ 知

$$\mu=E(X)=300\times 1/4=75,$$

$$\sigma^2=D(X)=300\times 1/4\times 3/4=225/4.$$

由契比雪夫不等式，有

$$P\{|X-E(X)|\leqslant 50\}\geqslant 1-\frac{225}{4}\Big/50^2=0.9775.$$

(2) 由 $X\sim B(1000,1/4)$ 知

$$\mu=E(X)=250,\quad \sigma^2=D(X)=375/2.$$

由契比雪夫不等式，有

$$P\{200\leqslant X\leqslant 300\}=P\{|X-250|\leqslant 50\}=P\{|X-\mu|\leqslant 50\}$$

$$\geqslant 1-\frac{375}{2}\Big/50^2=0.925.$$

例7 设在每次试验中，事件A发生的概率均为3/4. 用契比雪夫不等式估计，问：需要进行多少次独立重复试验，才能使事件发生的频率在0.74～0.76之间的概率至少为0.90?

解 设X为在n次独立重复试验中A发生的次数，确定$X\sim B(n,3/4)$. 以X/n表示在n次独立重复试验中A发生的概率，则

$$E\left(\frac{X}{n}\right)=\frac{1}{n}E(X)=\frac{1}{n}\times n\times\frac{3}{4}=\frac{3}{4},$$

$$D\left(\frac{X}{n}\right)=\frac{1}{n^2}D(X)=\frac{1}{n^2}n\times\frac{3}{4}\times\frac{1}{4}=\frac{3}{16n}.$$

由契比雪夫不等式，有

$$P\left\{0.74\leqslant\frac{X}{n}\leqslant 0.76\right\}$$

$$=P\left\{\left|\frac{X}{n}-0.75\right|\leqslant 0.01\right\}=P\left\{\left|\frac{X}{n}-E(X)\right|\leqslant 0.01\right\}$$

$$\geqslant 1-D\left(\frac{X}{n}\right)\Big/0.01^2=1-\frac{30000}{16n}.$$

所以，要使$P\left\{0.74\leqslant\frac{X}{n}\leqslant 0.76\right\}\geqslant 0.90$，即$1-\frac{30000}{16n}\geqslant 0.90$，知$n\geqslant 18750$，至少应进行18750次试验才能达到要求.

例8 设随机变量X的概率密度为

$$f(x)=\begin{cases}x^m e^{-x}/m!, & x\geqslant 0, m\text{ 为自然数},\\ 0, & \text{其它},\end{cases}$$

证明：
$$P\{0<x<2(m+1)\}\geqslant m/(m+1).$$

证 先计算X的数学期望与方差.

$$E(X)=\int_{-\infty}^{+\infty}xf(x)\mathrm{d}x=\frac{1}{m!}\int_{0}^{+\infty}x^{m+1}\mathrm{e}^{-x}\mathrm{d}x=\frac{1}{m!}\Gamma(m+2)$$

$$=\frac{1}{m!}(m+1)!=(m+1)\quad（利用\Gamma 函数性质），$$

$$E(X^2)=\int_{-\infty}^{+\infty}x^2f(x)\mathrm{d}x=\frac{1}{m!}\int_{0}^{+\infty}x^{m+2}\mathrm{e}^{-x}\mathrm{d}x=\frac{1}{m!}\Gamma(m+3)$$

$$=\frac{1}{m!}(m+2)!=(m+2)(m+1),$$

$$D(X)=E(X^2)-[E(X)]^2=(m+2)(m+1)-(m+1)^2=m+1.$$

由契比雪夫不等式，有

$$\begin{aligned}&P\{0<X<2(m+1)\}\\&=P\{-(m+1)<X-(m+1)<(m+1)\}\\&=P\{|X-(m+1)|<(m+1)\}\\&=P\{|X-E(X)|<m+1\}\geqslant 1-D(X)/(m+1)^2\\&=1-(m+1)/(m+1)^2=m/(m+1).\end{aligned}$$

本题和例 4、例 6 中都有一个小小的技巧，就是将所求概率式化为契比雪夫不等式形式. 往往是在求出期望值后，照顾期望值而得出 ε 值，再利用契比雪夫不等式的结论.

例9 已知正常男性成人每毫升血液中平均白细胞数是7300，标准差是700. 利用契比雪夫不等式估计男性成人每毫升血液中含白细胞数在 5200 至 9400 之间的概率 p.

解 确定每毫升血液中白细胞数为随机变量 X，由题设

$$E(X)=7300,\quad D(X)=700^2.$$

由契比雪夫不等式，将概率式转化为不等式，得

$$\begin{aligned}P\{5200<X<9400\}&=P\{|X-7300|<2100\}\\&\geqslant 1-700^2/2100^2=8/9.\end{aligned}$$

例 10 已知 $D(X)=0$，证明：$P\{X=E(X)\}=1$.

证 $P\{X=E(X)\}=P\{X-E(X)=0\}$

$$=1-P\{X-E(X)\neq 0\},$$

而 $\{X-E(X)\neq 0\}=\{|X-E(X)|\neq 0\}$

$$=\bigcup_{n=1}^{\infty}\{|X-E(X)|\geqslant 1/n\}.$$

这一步转化十分重要，由此可利用契比雪夫不等式，因为

$$P\{|X-E(X)|\neq 0\}=\sum_{n=1}^{\infty}P\left\{|X-E(X)|\geqslant\frac{1}{n}\right\},$$

而 $$P\left\{|X-E(X)|\geqslant\frac{1}{n}\right\}\leqslant D(X)\Big/\frac{1}{n}=0,$$

所以 $$P\{|X-E(X)|\neq 0\}=0,\ n=1,2,\cdots,$$

于是 $$P\{|X-E(X)|=0\}=1-0=1.$$

二、大数定律及应用

大数定律的应用，关键是要找到一个随机变量序列，根据计算出的 $E(X_i)$ 和 $D(X_i)$，确定定理条件是否满足，然后依据大数定律解决问题. 常用的方法和技巧是：求 $E(X_i)$ 和 $D(X_i)$ 要用到计算数学期望和方差的技巧，利用契比雪夫不等式，利用有关定理与公式（如斯特林格公式），利用反证法，等等. 一定要根据具体情况灵活运用方法和技巧.

例 11 设有随机变量序列 $\{X_k\}$ 且相互独立，其分布律为

X_k	$-\sqrt{2}$	0	$\sqrt{2}$
p_k	1/4	1/2	1/4

问：可否对此随机变量序列使用大数定律？

解 已知 X_k 相互独立，$E(X_k)=0$，$E(X_k^2)=1$，又

$$D(X_k)=E(X_k^2)-[E(X_k)]^2=1-0=1,$$

所以，满足契比雪夫大数定理条件，可使用大数定律.

例 12 设随机变量序列 $X_1,X_2,\cdots,X_n$ 相互独立且同分布，X_i 的分布律为

X_i	$-ia$	0	ia
p_k	$1/(2i^2)$	$1-1/i^2$	$1/(2i^2)$

证明：$\lim\limits_{n\to\infty}P\left\{\left|\frac{1}{n}\sum_{i=1}^{n}X_i\right|>\varepsilon\right\}=0.$

证 先验证是否满足大数定律条件. 因为

$$E(X_i)=0,\quad E(X_i^2)=2i^2a^2\times 1/(2i^2)=a^2,$$

所以
$$E\left(\frac{1}{n}\sum_{i=1}^{n}X_i\right)=\frac{1}{n}\sum_{i=1}^{n}E(X_i)=0,$$

$$D\left(\frac{1}{n}\sum_{i=1}^{n}X_i\right)=\frac{1}{n^2}\sum_{i=1}^{n}D(X_i)=\frac{a^2}{n},$$

满足契比雪夫定理条件，有

$$\lim_{n\to\infty}P\left\{\left|\frac{1}{n}\sum_{i=1}^{n}X_i-0\right|<\varepsilon\right\}=1,$$

即
$$\lim_{n\to\infty}P\left\{\left|\frac{1}{n}\sum_{i=1}^{n}X_i\right|>\varepsilon\right\}=0.$$

也可以由契比雪夫不等式得出

$$P\left\{\left|\frac{1}{n}\sum_{i=1}^{n}X_i-0\right|>\varepsilon\right\}\leqslant D\left(\frac{1}{n}\sum_{i=1}^{n}X_i\right)\Big/\varepsilon^2=\frac{a^2}{n\varepsilon^2},$$

故
$$\lim_{n\to\infty}P\left\{\left|\frac{1}{n}\sum_{i=1}^{n}X_i\right|>\varepsilon\right\}=0.$$

例 13 已知独立随机变量序列 $X_1,X_2,\cdots,X_n$ 具有同一分布

$$F(x)=1/2+1/\pi\cdot\arctan(x/a),$$

问：是否可以适用辛钦大数定律？

解 因为 $f(x)=a/[\pi(a^2+x^2)]$，所以

$$E(X)=\int_{-\infty}^{+\infty}|x|\frac{a}{\pi(a^2+x^2)}\mathrm{d}x=\frac{2a}{\pi}\int_0^{+\infty}\frac{ax}{a^2+x^2}\mathrm{d}x=\infty,$$

而辛钦大数定律要求 $E(X)$ 存在，故该随机变量序列不适用辛钦大数定律.

例 14 设 $X_1,X_2,\cdots,X_n,\cdots$ 为相互独立且同分布的随机变量序列，服从 $U(0,1)$，证明：

$$\left(\prod_{k=1}^{n}X_k\right)^{1/n}\xrightarrow{p}C,\quad n\to\infty.$$

其中 C 为常数，并求出 C 的值.

证 令 $Y_n=\ln X_n$，则 Y_n $(n\geqslant 1)$ 为相互独立且同分布的随机变量序列，有

$$E(Y_n)=\int_0^1 \ln x \mathrm{d}x = x\ln x\Big|_0^1-\int_0^1 \mathrm{d}x=0-1=-1,$$

满足辛钦大数定律,则

$$\frac{1}{n}\sum_{k=1}^{n}Y_k \xrightarrow{p} -1,\quad n\to\infty.$$

故 $$\left(\prod_{k=1}^{n}X_k\right)^{1/n}=\exp\left(\frac{1}{n}\sum_{k=1}^{n}Y_k\right)\xrightarrow{p}\mathrm{e}^{-1},\quad n\to\infty,$$

于是 $$C=\mathrm{e}^{-1}.$$

例 15 证明马尔柯夫大数定理:如果随机变量序列 $X_1,X_2,\cdots,X_n,\cdots$ 满足 $\lim\limits_{n\to\infty}\frac{1}{n^2}D\left(\sum\limits_{k=1}^{n}X_k\right)=0$,则对任意 $\varepsilon>0$,有

$$\lim_{n\to\infty}P\left\{\left|\frac{1}{n}\sum_{k=1}^{n}X_k-\frac{1}{n}\sum_{k=1}^{n}E(X_k)\right|<\varepsilon\right\}=1.$$

证 先求 $\frac{1}{n}\sum\limits_{k=1}^{n}X_k$ 的数学期望与方差.因为

$$E\left(\frac{1}{n}\sum_{k=1}^{n}X_k\right)=\frac{1}{n}\sum_{k=1}^{n}E(X_k),\quad D\left(\frac{1}{n}\sum_{k=1}^{n}X_k\right)=\frac{1}{n^2}\sum_{k=1}^{n}D(X_k),$$

由契比雪夫不等式,有

$$P\left\{\left|\frac{1}{n}\sum_{k=1}^{n}X_k-\frac{1}{n}\sum_{k=1}^{n}E(X_k)\right|<\varepsilon\right\}\geqslant 1-D\left(\sum_{k=1}^{n}X_k\right)\Big/(n^2\varepsilon^2).$$

由题给条件,立即可得

$$\lim_{n\to\infty}P\left\{\left|\frac{1}{n}\sum_{k=1}^{n}X_k-\frac{1}{n}\sum_{k=1}^{n}E(X_k)\right|<\varepsilon\right\}=1.$$

例 16 设 $X_1,X_2,\cdots,X_n,\cdots$ 同分布,当 $|k-i|\geqslant 2$ 时,X_k 与 X_i 相互独立,方差 $D(X_i)$ 存在,证明:

$$\lim_{n\to\infty}P\left\{\left|\frac{1}{n}\sum_{i=1}^{n}X_i-\frac{1}{n}\sum_{i=1}^{n}E(X_i)\right|<\varepsilon\right\}=1.$$

证 为简单计,不妨设

$$E(X_i)=0,\quad E(X_i^2)=D(X_i)=1,\ i=1,2,\cdots,$$

于是

$$D\Big(\sum_{i=1}^{n}X_i\Big)=E\Big(\sum_{i=1}^{n}X_i\Big)^2=\sum_{i=1}^{n}E(X_i^2)+2\sum_{1\leqslant i<k\leqslant n}E(X_iX_i)$$
$$=n+2[E(X_1X_2)+E(X_2X_3)+\cdots+E(X_{n-1}X_n)]$$
$$\leqslant n+2n=3n.$$

这里利用了以下条件：当$i\neq k$时，

$$|E(X_iX_k)|\leqslant\sqrt{E(X_i^2)+E(X_k^2)}=1,$$

从而 $$D\Big(\frac{1}{n}\sum_{i=1}^{n}X_i\Big)=\frac{1}{n^2}D\Big(\sum_{i=1}^{n}X_i\Big)\leqslant\frac{3}{n}\longrightarrow 0\ (n\to\infty),$$

即满足马尔柯夫定理条件，所以

$$\lim_{n\to\infty}P\left\{\left|\frac{1}{n}\sum_{i=1}^{n}X_i-\frac{1}{n}\sum_{i=1}^{n}E(X_i)\right|<\varepsilon\right\}=1.$$

例 17 设$X_1,X_2,\cdots,X_n,\cdots$是相互独立的随机变量序列，且 $P\{X_n=\pm\sqrt{\ln n}\}=\frac{1}{2}$，$n=1,2,\cdots$，验证：$\{X_n\}$服从大数定律.

证一 $E(X_i)=0$， $D(X_i)=E(X_i^2)=\ln i$，且有

$$D\Big(\sum_{i=1}^{n}X_i\Big)=\sum_{i=1}^{n}D(X_i)=\sum_{i=1}^{n}\ln i\leqslant n\ln n,$$

所以 $$\frac{1}{n^2}D\Big(\sum_{i=1}^{n}X_i\Big)\leqslant\frac{\ln n}{n}\longrightarrow 0\ (n\to\infty).$$

于是，由马尔柯夫定理知，随机变量序列$\{X_n\}$服从大数定律.

证二 由斯特林格公式$n!=n^{n+1/2}e^{-n\sqrt{2\pi}}$，有

$$\frac{1}{n^2}D\Big(\sum_{i=1}^{n}X_i\Big)=\frac{1}{n^2}\sum_{i=1}^{n}\ln i=\frac{1}{n!}\ln(n!)$$
$$\approx\frac{1}{n^2}\left[\Big(n+\frac{1}{2}\Big)\ln n-n+\ln(n\sqrt{2\pi})\right]\longrightarrow 0,$$

于是，由马尔柯夫定理知，随机变量序列$\{X_k\}$服从大数定律.

例 18 设$\{X_n\}$为相互独立的随机变量序列，且

$$P\{X_n=\pm 2^n\}=1/2^{2n+1},$$
$$P\{X_n=0\}=1-1/2^{2n},\ n=1,2,\cdots,$$

证明：$\{X_n\}$服从大数定律.

证 $E(X_n)=2^n\frac{1}{2^{2n+1}}-2^n\frac{1}{2^{2n+1}}+0\times\left(1-\frac{1}{2^{2n}}\right)=0,$

$$D(X_n)=2^{2n}\frac{1}{2^{2n+1}}+2^{2n}\frac{1}{2^{2n+1}}+0\times\left(1-\frac{1}{2^{2n}}\right)=1.$$

令$Y_n=\frac{1}{n}\sum_{i=1}^{n}X_i$, $n=1,2,\cdots$,则$E(Y_n)=0$,$D(Y_n)=\frac{1}{n}$. 于是,对任意$\varepsilon>0$,有契比雪夫不等式

$$P\{|Y_n-E(Y_n)|<\varepsilon\}\geqslant 1-\frac{1}{n\varepsilon^2},$$

即有
$$\lim_{n\to\infty}P\{|Y_n-E(Y_n)|<\varepsilon\}=1,$$

所以,随机变量序列$\{X_n\}$服从大数定律.

例19 设$X_1,X_2,\cdots,X_n,\cdots$是相互独立且同分布的随机变量,$X_i\sim U(a,b)$,$f(x)$是$[a,b]$上的连续函数. 证明:

$$p-\lim_{n\to\infty}\frac{b-a}{n}\sum_{i=1}^{n}f(X_i)=\int_a^b f(x)\mathrm{d}x.$$

证 由$X_1,X_2,\cdots,X_n,\cdots$相互独立知,$f(X_1),f(X_2),\cdots,f(X_n),\cdots$也相互独立. 因为$X_i$的密度函数为

$$g(x)=\begin{cases}1/(b-a), & a\leqslant x\leqslant b,\\ 0, & \text{其它},\end{cases}$$

所以 $E[(b-a)f(X_i)]=\int_{-\infty}^{+\infty}(b-a)f(x)g(x)\mathrm{d}x=\int_a^b f(x)\mathrm{d}x.$

由辛钦大数定律,有

$$p-\lim_{n\to\infty}\frac{b-a}{n}\sum_{i=1}^{n}f(X_i)=E[(b-a)f(X_i)]=\int_a^b f(x)\mathrm{d}x.$$

若把本题看成积分和的极限,则它提供了一种定积分计算的方法.

例20 以某种仪器测量已知量A时,设n次独立得到的测量值为$x_1,x_2,\cdots,x_n$. 如果仪器无系统误差,问:当n充分大时,是否可以取$\frac{1}{n}\sum_{i=1}^{n}(x_i-A)^2$作为仪器测量误差方差的近似值?

解 若把x_i视为n个相互独立同分布随机变量X_i $(i=1,2,$

$\cdots,n$)的观察值，则$E(X_i)=\mu,D(X_i)=\sigma^2$ $(i=1,2,\cdots,n)$. 仪器第i次测量的误差 X_i-A 的数学期望和

$$E(X_i-A)=\mu-A,$$

方差分别为

$$D(X_i-A)=\sigma^2.$$

设$Y_i=(X_i-A)^2$, $i=1,2,\cdots,n$，则Y_i 也相互独立，服从同一分布. 在仪器无系统误差时，有$E(X_i-A)=0$，即$\mu=A$，于是

$$\begin{aligned}E(Y_i)&=E[(X_i-A)^2]=E\{[X_i-E(X_i)]^2\}\\&=D(X_i)=\sigma^2,\ i=1,2,\cdots,n.\end{aligned}$$

由契比雪夫定理的特殊情形，可得

$$\lim_{n\to\infty}P\left\{\left|\frac{1}{n}\sum_{i=1}^{n}Y_i-\sigma^2\right|<\varepsilon\right\}=1,$$

即

$$\lim_{n\to\infty}P\left\{\left|\frac{1}{n}\sum_{i=1}^{n}(X_i-A)^2-\sigma^2\right|<\varepsilon\right\}=1.$$

从而确定，当$n\to\infty$时，随机变量$\frac{1}{n}\sum\limits_{i=1}^{n}(X_i-A)^2$ 依概率收敛于σ^2. 即当n 充分大时，可以取$\frac{1}{n}\sum\limits_{i=1}^{n}(X_i-A)^2$ 作为仪器测量误差的方差的近似值.

第二节　中心极限定理

主要内容

1. 列维-林德伯格定理(同分布的中心极限定理)

设$X_1,X_2,\cdots,X_n,\cdots$是相互独立且同分布的随机变量序列，有有限的数学期望与方差，$E(X_i)=\mu,D(X_i)=\sigma^2\neq0$ $(i=1,2,\cdots)$，则对任意实数x，随机变量

$$Y_n=\frac{\sum_{i=1}^{n}(X_i-\mu)}{\sqrt{n}\,\sigma}=\frac{\sum_{i=1}^{n}X_i-n\mu}{\sqrt{n}\,\sigma}$$

的分布函数 $F_n(x)$ 满足

$$\lim_{n\to\infty}F_n(x)=\lim P\{Y_n\leqslant x\}=\int_{-\infty}^{x}\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-t^2/2}\mathrm{d}t.$$

2. 德莫弗-拉普拉斯定理

设随机变量 η_n $(n=1,2,\cdots)$ 服从参数为 n,p $(0<p<1)$ 的二项分布，则对于任意的 x，恒有

$$\lim_{n\to\infty}P\left\{\frac{\eta_n-np}{\sqrt{np(1-p)}}\leqslant x\right\}=\int_{-\infty}^{x}\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-t^2/2}\mathrm{d}t.$$

3. 李雅普诺夫定理

设 $X_1,X_2,\cdots,X_n,\cdots$ 是相互独立且不同分布的随机变量，它们分别有数学期望和方差

$$E(X_i)=\mu_i,\quad D(X_i)=\sigma_i^2\neq 0,\ i=1,2,\cdots.$$

记 $B_n^2=\sum_{i=1}^{n}\sigma_i^2$，若存在正数 δ，使得当 $n\to\infty$ 时，有

$$\frac{1}{B_n^{2+\delta}}\sum_{i=1}^{n}E\{|X_i-\mu_i|^{2+\delta}\}\longrightarrow 0,$$

则随机变量

$$Z_n=\frac{\sum_{i=1}^{n}X_i-E\left(\sum_{i=1}^{n}X_i\right)}{\sqrt{D\left(\sum_{i=1}^{n}X_i\right)}}=\frac{\sum_{i=1}^{n}X_i-\sum_{i=1}^{n}\mu_i}{B_n}$$

的分布函数 $F_n(x)$ 对于任意的 x，满足

$$\lim_{n\to\infty}F_n(x)=\lim_{n\to\infty}\left(\frac{\sum_{i=1}^{n}X_i-\sum_{i=1}^{n}\mu_i}{B_n}\leqslant x\right)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}\mathrm{e}^{-t^2/2}\mathrm{d}t.$$

疑难解析

1. 中心极限定理有什么实际意义?

答 正态分布是概率论中三个重要分布之一,它是现实生活和科学技术中使用最多的一种分布,也是数理统计的重要假设.许多随机变量本身并不属于正态分布,但它们的共同作用下形成的随机变量的极限分布是正态分布.它们的概率如何计算是一个很重要的问题.中心极限定理阐明了,在什么条件下原本不属于正态分布的一些随机变量其总和分布渐近服从正态分布.

2. 大数定律与中心极限定理有什么异同?

答 大数定律与中心极限定理都是通过极限理论来研究概率问题,研究对象都是随机变量序列,解决的都是概率论中的基本问题,因而大数定律与中心极限定理在概率论中的意义十分重要.

它们的不同在于:大数定律给出的是当$n\to\infty$时随机变量序列的函数(平均值或概率)的极限;而中心极限定理则告诉我们,随机变量序列总和的分布近似正态分布,总和的标准化随机变量服从渐近标准正态分布,而不论随机变量序列服从何种分布.这个问题是近两个世纪来概率论研究的中心问题,所以称为中心极限定理.

方法、技巧与典型例题分析

应用中心极限定理的关键是,由所给条件构造一个相互独立同分布的随机变量序列,使具有有限的数学期望与方差,然后建立一个标准化随机变量,即可应用中心极限定理.

用中心极限定理讨论的第一个问题是:在概率确定的条件下求样本数n.解题方法是确定随机变量序列,建立标准化随机变量.建立已知概率与标准正态分布的关系式,寻找n的表达式,通过正态分布表得出n的关系式即可求出n.技巧主要表现在求随机变量

的数学期望和方差上，一般技巧是通过分离组合或者随机变量函数的形式来给出和的数学期望与方差，如例1、例3、例7、例9等.

第二个问题是：计算总和在某一区间内的概率. 其一般方法与第一个问题相同，只是具体操作时后者只需计算两个标准正态分布函数值之差就可以了. 没有什么特别的技巧，只要按照一般的方法去做即可，如例2、例5、例8等.

下面通过例题来了解解题的方法和技巧.

例1 在抽样检查某种产品质量时，如果发现次品多于10个，则拒绝接受这批产品. 设产品的次品率为10%，问：至少应抽取多少个产品进行检查，才能保证拒绝接受这批产品的概率达到0.9?

解 设n为应抽取的产品数，Y为其中的次品数，记

$$X_i=\begin{cases}1, & \text{第}\,i\,\text{次检查时为次品},\\ 0, & \text{第}\,i\,\text{次检查时为正品},\end{cases}$$

则

$$Y=\sum_{i=1}^{n}X_i,\quad E(X_i)=0.1,$$

$$D(X_i)=0.1\times(1-0.1)=0.09.$$

由德莫弗-拉普拉斯定理，得

$$\begin{aligned}&P\{10<Y\leqslant n\}\\&=P\left\{\frac{10-n\times0.1}{\sqrt{n\times0.1\times0.9}}<\frac{Y-n\times0.1}{\sqrt{n\times0.1\times0.9}}\leqslant\frac{n-n\times0.1}{\sqrt{n\times0.1\times0.9}}\right\}\\&=\Phi(3\sqrt{n})-\Phi\left(\frac{10-0.1n}{0.3\sqrt{n}}\right)\approx1-\Phi\left(\frac{10-0.1n}{0.3\sqrt{n}}\right).\end{aligned}$$

因为

$$1-\Phi\left(\frac{10-0.1n}{0.3\sqrt{n}}\right)=0.9,$$

所以

$$\frac{10-0.1n}{0.3\sqrt{n}}=1.28\Longrightarrow n=147,$$

即至少应抽取147个产品检查才能达到目的.

例2 某车间有150台同类型的机器，每台机器出现故障的概率都是0.02. 设各台机器的工作是相互独立的，求机器出现故障的

台数不少于 2 的概率.

解 设机器出现故障的台数为 X, $X\sim B(150,0.02)$, 则 $E(X)=3$, $D(X)=2.94$, $\sqrt{D(X)}=1.715$, 由中心极限定理, 有

$$P\{X\geqslant 2\}=1-P\{X\leqslant 1\}=1-P\left\{\frac{X-3}{1.715}\leqslant\frac{2-3}{1.715}\right\}$$
$$=1-\Phi(-0.5832)=0.7201.$$

例 3 某单位有 1000 人独立地参加防空演习, 设每个人能按时进入掩体的概率为 0.9, 以 0.95 的概率估计:

(1) 在一次演习中至少有多少人能进入掩体?

(2) 在一次演习中至多有多少人能进入掩体?

解 用 X_i $(i=1,2,\cdots,1000)$ 表示第 i 人能按时进入掩体, 令

$$S_m=X_1+X_2+\cdots+X_m.$$

(1) 设至少有 m 人能进入掩体, 使得 $P\{m\leqslant S_m\leqslant 1000\}\geqslant 0.95$, 因为

$$\{m\leqslant S_m\}=\left\{\frac{m-1000\times 0.9}{\sqrt{1000\times 0.9\times 0.1}}\leqslant\frac{S_m-1000\times 0.9}{\sqrt{1000\times 0.9\times 0.1}}\right\},$$

令 $\dfrac{S_m-900}{\sqrt{90}}=Y$, 则 $Y\sim N(0,1)$, 由中心极限定理, 有

$$P\{m\leqslant S_m\}=P\left\{Y\geqslant\frac{m-900}{\sqrt{90}}\right\}=1-P\left\{Y<\frac{m-900}{\sqrt{90}}\right\}$$
$$=1-\Phi\left(\frac{m-900}{\sqrt{90}}\right)=0.95.$$

查正态分布表知 $\dfrac{m-900}{\sqrt{90}}=-1.65$, 得 $m=884.35$, 即至少有 884 人能进入掩体.

(2) 用类似方法可知, 至多有 916 人能进入掩体 (请读者一试).

例 4 设随机变量序列 $\{X_n\}$ 相互独立, 在 $[-n,n]$ 上 X_n 服从均匀分布, 问: 能否对 $\{X_n\}$ 使用中心极限定理?

解 能. 因为 $X_n\sim U(-n,n)$, $n=1,2,\cdots$, 所以 $E(X_n)=0$,

$D(X_n)=(n+n)^2/12=n^2/3$. 满足列维-林德伯格定理条件，有有限的数学期望与方差，可对$\{X_n\}$使用中心极限定理.

例5 某电站对一万个用户供电. 设用电高峰时每户用电的概率为0.9，利用中心极限定理，计算：

(1) 同时用电户数在9030户以上的概率；

(2) 若每户用电200 W，电站至少应具有多大发电量，才能以0.95的概率保证供电.

解 以X记用电高峰时同时用电的户数.

(1) 所求概率为$P\{X>9030\}$. 因为

$$X\sim B(10000,0.9),\quad E(X)=9000,\quad D(X)=900,$$

于是 $$P\{X>9030\}=P\left\{\frac{10000-9000}{\sqrt{900}}\geqslant\frac{9030-9000}{\sqrt{900}}\right\}$$

$$=\Phi\left(\frac{100}{3}\right)-\Phi(1)\approx 1-0.8413=0.1587.$$

(2) 设电站的发电量(单位:W)至少为x才能以0.95的概率保证供电，则因为要

$$P\{200X\leqslant x\}=P\left\{X\leqslant\frac{x}{200}\right\}=P\left\{\frac{X-9000}{30}\leqslant\frac{x/200-9000}{30}\right\}$$

$$=\Phi\left(\frac{x-1800000}{6000}\right)-0\geqslant 0.95,$$

所以$\frac{x-1800000}{6000}\geqslant 1.65$，得$x\geqslant 1809900$，即电站具有1809900 W发电量，才能以0.95的概率保证供电.

例6 现从某厂生产的一批同型号电子元件中抽取395件，由于次品率未知，需要通过次品的相对频率来估计，这时估计的可靠性大于95%. (1) 求绝对误差ε；(2) 如果样品中有十分之一是次品，应对p怎样估计？

解 以β记元件的可靠度.

(1) $\beta=P\left\{\left|\frac{\eta_n}{n}-p\right|<\varepsilon\right\}\approx 2\Phi[\varepsilon\sqrt{n/(pq)}]-1$，即

$$\Phi[\varepsilon\sqrt{n/(pq)}]\geqslant(1+\beta)/2.$$

由 $\beta=0.95$ 得 $(1+\beta)/2=0.975$，查正态分布表知

$$\varepsilon\sqrt{n/(pq)}=1.96 \Longrightarrow \varepsilon=1.96\sqrt{pq/n}.$$

因为 $p+q=1$，所以 $pq\leqslant 1/4$，$n=395$，故

$$\varepsilon\leqslant 1.96/(2\sqrt{395})=0.05.$$

(2) 由 $P\left\{\left|\frac{\eta_{395}}{395}-p\right|<\varepsilon\right\}\geqslant\beta$ 知

$$\frac{\eta_{395}}{395}-\varepsilon<p<\frac{\eta_{395}}{395}+\varepsilon,\quad \beta=0.95,$$

而次品的频率(由题设) $\frac{\eta_{395}}{395}=\frac{1}{10}$，从而得

$$0.10<p<0.15.$$

例 7 设一条自动生产线的产品合格率是 0.8. 要使一批产品的合格率在 76% 与 84% 之间的概率不小于 90%，问：这批产品至少要生产多少件？

解 设至少要生产 m 件产品，并以 X 记 m 件产品中合格品的件数，则 $X\sim B(m,0.8)$. 现在要确定 m，使满足概率不等式

$$P\{0.76<X/m<0.84\}\leqslant 0.90.$$

(1) 若用契比雪夫不等式估计，有

$$\begin{aligned}&P\{0.76<X/m<0.84\}\\&=P\{|X-0.8m|<0.04m\}\\&\geqslant 1-(0.8m\times 0.2)/(0.04m)^2=1-100/m.\end{aligned}$$

由 $1-100/m\geqslant 0.90$，得 $m\geqslant 1000$，即至少要生产 1000 件产品才能保证合格品的概率满足要求.

(2) 用德莫弗-拉普拉斯定理估计，可知当 n 比较大时，X 近似服从正态分布 $N(0.8m,0.16m)$，于是

$$\begin{aligned}P\{0.76<X/m<0.84\}&=P\left\{\left|\frac{X-0.8m}{0.4\sqrt{m}}\right|<\frac{0.04m}{0.4\sqrt{m}}\right\}\\&\approx 2\Phi(0.1\sqrt{m})-1\geqslant 0.90.\end{aligned}$$

查正态分布表知，$0.1\sqrt{m}=1.65$，解得 $m\geqslant 268.96$，故取

$$m=269.$$

将题(1)与题(2)比较可知,由中心极限定理估计的m远比由契比雪夫不等式估计的m精确得多.

例 8 某地进行的抽样调查结果显示,考生的外语成绩(百分制)近似服从正态分布,平均成绩为72分,96分以上的占考生总数的2.3%.试求考生的外语成绩在60分至84分之间的概率.

解 以X记考生的外语成绩,则$X\sim N(72,\sigma^2)$.又由题设知

$$P\{X\geqslant 96\}=P\left\{\frac{X-72}{\sigma}\geqslant\frac{96-72}{\sigma}\right\}=1-\Phi\left(\frac{24}{\sigma}\right)=0.023,$$

即$\Phi\left(\frac{24}{\sigma}\right)=0.977$.查正态分布表知,$24/\sigma=2$,解得$\sigma=12$.所以,$X\sim N(72,12^2)$.于是

$$\begin{aligned}P\{60\leqslant X\leqslant 84\}&=P\left\{\frac{60-72}{12}\leqslant X\leqslant\frac{84-72}{12}\right\}\\&=\Phi(1)-\Phi(-1)=2\Phi(1)-1=0.682.\end{aligned}$$

例 9 某厂生产的产品次品率为$p=0.1$.为了确保销售,该厂向顾客承诺每盒中有100只以上正品的概率达到95%,问:该厂需要在一盒中装多少只产品?

解 设每盒中装m只产品,合格品数$X\sim B(m,0.9)$,$E(X)=0.9m$,$D(X)=0.09m$,则

$$P\{X>100\}=1-P\{X\leqslant 100\}=1-\Phi\left(\frac{100-0.9m}{0.3\sqrt{m}}\right)=0.95,$$

所以$\frac{100-0.9m}{0.3\sqrt{m}}=-1.65$,解得$m=117$,即每盒至少要装117只产品才能以95%的概率保证一盒内有100只正品.

例 10 某药厂对自己的一种药品的广告宣称,这种药品对疾病的治愈率为0.8.卫生部门任意抽查了100个服用此药的人,如果其中有多于75人治愈,就认为宣称是真实的,否则就是虚假的.

(1) 若此药的实际治愈率为0.8,求接受这一宣称的概率;

(2) 若此药的实际治愈率为0.7,求接受这一宣称的概率.

解 (1) 设100人中治愈的病人数为X，$X\sim B(100,0.8)$，$E(X)=80$，$D(X)=16$，依德莫弗-拉普拉斯定理，$(X-80)/4\sim N(0,1)$，于是

$$\begin{aligned}P\{75<X\leqslant 100\}&=P\left\{\frac{75-80}{4}<\frac{X-80}{4}\leqslant\frac{100-80}{4}\right\}\\&=\Phi(5)-\Phi(-1.25)\\&=\Phi(5)-1+\Phi(1.25)=0.8944,\end{aligned}$$

即接受这一宣称的概率为0.8944.

(2) 此时，$X\sim B(100,0.7)$，$D(X)=21$，$E(X)=70$. 依德莫弗-拉普拉斯定理，$(X-70)/\sqrt{21}\sim N(0,1)$，故

$$\begin{aligned}P\{75<X\leqslant 100\}&=P\left\{\frac{75-70}{\sqrt{21}}<\frac{X-70}{\sqrt{21}}\leqslant\frac{100-70}{\sqrt{21}}\right\}\\&=\Phi(6.547)-\Phi(1.091)=0.1379.\end{aligned}$$

例11 某厂生产的产品平均寿命为2000 h，标准差为250 h. 进行技术改造后，平均寿命提高到2250 h，标准差不变. 为了确认这一成果，检验的方法是：任意选取若干件产品进行测试，若产品平均寿命超过2200 h，就确认技术改造成功. 要使检验通过的概率超过0.997，至少应检验多少件产品？

解 设应检验n件产品. 以X_i记第i件产品的使用寿命，则

$$\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i,\quad E(X_i)=2250,\quad D(X_i)=250^2.$$

由列维-林德伯格定理，有

$$\begin{aligned}&P\left\{\frac{1}{n}\sum_{i=1}^{n}X_i>2200\right\}\\&=P\left\{\left[\sum_{i=1}^{n}-nE(X_i)\right]\Big/(\sqrt{n}\,\sigma)>\left[2200-E(X_i)\sqrt{n}\right]/250\right\}\\&=P\left\{\left[\sum_{i=1}^{n}X_i-nE(X_i)\right]\Big/(\sqrt{n}\,\sigma)>-\sqrt{n}/5\right\}\\&=1-\Phi(-\sqrt{n}/5)=\Phi(\sqrt{n}/5)\geqslant 0.997.\end{aligned}$$

查正态分布表知，$\sqrt{n}/5 \geqslant 2.75$，解得 $n \geqslant 189.0625$，故取 $n=190$.

例 12 某地有甲、乙两个电影院竞争当地的 1000 名观众，观众选择电影院是相互独立的和随机的，问：每个电影院至少应设有多少个座位，才能保证观众因缺少座位而离去的概率小于 1%？

解 不妨设甲、乙两影院是对称的，故只需讨论甲影院的情形. 设

$$X_k = \begin{cases} 1, & \text{第 } k \text{ 个观众选择甲影院}, \\ 0, & \text{其它}, \end{cases}$$

则甲影院的观众人数总数 $X = \sum_{k=1}^{1000} X_k$，而

$$E(X_k) = \frac{1}{2}, \quad D(X_k) = E(X_k^2) - [E(X_k)]^2 = \frac{1}{2} - \left(\frac{1}{2}\right)^2 = \frac{1}{4}$$

（因为 $P\{X_k = 1\} = 1/2$）. 又

$$n = 1000, \quad nE(X_k) = n\mu = 500, \quad n\sigma^2 = 250.$$

依列维-林德伯格定理，有 $(X-500)/(5\sqrt{10}) \sim N(0,1)$. 于是，若应设座位数为 M，则

$$P\{X \leqslant M\} = P\{(X-500)/(5\sqrt{10}) \leqslant (M-500)/(5\sqrt{10})\}$$
$$= \Phi[(M-500)/(5\sqrt{10})] \geqslant 0.99.$$

查正态分布表知，$(M-500)/(5\sqrt{10}) \geqslant 2.33$，$M$ 应取 537，即每个电影院至少应设 537 个座位，才能符合要求.

例 13 设 $X_1, X_2, \cdots, X_n$ 相互独立，且都服从泊松分布 $\pi(1)$. 令 $Y_n = \sum_{i=1}^{n} X_i$，证明：当 $n \to \infty$ 时，$(Y_n - n)/n$ 的极限分布是标准正态分布.

证 因为 $E(X_i) = 1$，$D(X_i) = 1$，$i = 1, 2, \cdots, n$，

所以 $E(Y_n) = E\left(\sum_{i=1}^{n} X_i\right) = n$，$D(Y_n) = D\left(\sum_{i=1}^{n} X_i\right) = n$.

由列维-林德伯格定理，有

$$\frac{Y-E(Y_n)}{\sqrt{nD(Y_n)}}=\frac{Y-n}{n}\longrightarrow N(0,1).$$

例 14 证明：当$n\to\infty$时，有

$$\left(1+n+\frac{1}{2!}n^2+\cdots+\frac{1}{n!}n^n\right)\mathrm{e}^{-n}\longrightarrow\frac{1}{2}.$$

证 设X_n（$n\geqslant 1$）是相互独立且同分布的随机变量序列，X_n服从参数为1的泊松分布，则对每个n，S_n服从参数为n的泊松分布. 依列维-林德伯格定理，有

$$(S_n-n)/\sqrt{n}\stackrel{\cdot}{\sim}N(0,1),$$

所以，由标准正态分布的对称性，得

$$\begin{aligned}\frac{1}{2}&=\lim_{n\to\infty}P\{(S_n-n)/\sqrt{n}\leqslant 0\}=\lim_{n\to\infty}\{S_n\leqslant n\}\\&=\lim_{n\to\infty}\left(1+n+\frac{1}{2!}n^2+\cdots+\frac{1}{n!}n^n\right)\mathrm{e}^{-n}.\end{aligned}$$

硕士研究生入学试题分析

一、本章考试要求

1. 了解契比雪夫不等式.

2. 了解契比雪夫大数定律、伯努利大数定律和辛钦大数定律（独立同分布随机变量的大数定律）成立的条件及结论.

3. 了解列维-林德伯格定理（独立同分布的中心极限定理）和德莫弗-拉普拉斯定理（二项分布以正态分布为极限分布）的应用条件和结论，并会用相关定理近似计算有关随机事件的概率.

二、本章重点内容

契比雪夫不等式及其应用，列维-林德伯格定理及其应用.

大数定律与中心极限定理.

1. 设$X_1,X_2,\cdots,X_n,\cdots$为相互独立且同分布的随机变量序

列，并均服从参数为λ ($\lambda>1$)的指数分布，记$\Phi(x)$为标准正态分布函数，则(　　).

(A) $\lim\limits_{n\to\infty}P\left\{\dfrac{\sum\limits_{i=1}^{n}X_i-n\lambda}{\lambda\sqrt{n}}\leqslant x\right\}=\Phi(x)$;

(B) $\lim\limits_{n\to\infty}P\left\{\dfrac{\sum\limits_{i=1}^{n}X_i-n\lambda}{\sqrt{n\lambda}}\leqslant x\right\}=\Phi(x)$;

(C) $\lim\limits_{n\to\infty}P\left\{\dfrac{\lambda\sum\limits_{i=1}^{n}X_i-n}{\sqrt{n}}\leqslant x\right\}=\Phi(x)$;

(D) $\lim\limits_{n\to\infty}P\left\{\dfrac{\sum\limits_{i=1}^{n}X_i-\lambda}{\sqrt{n\lambda}}\leqslant x\right\}=\Phi(x)$.

(2005 年四)

解　选(C). 因为$E(X_i)=1/\lambda, D(X_i)=1/\lambda^2$，所以

$$E\left(\sum_{i=1}^{n}X_i\right)=n/\lambda,\quad D\left(\sum_{i=1}^{n}X_i\right)=n/\lambda^2.$$

2. 设随机变量X的数学期望$E(X)=\mu$，方差$D(X)=\sigma^2$，则由契比雪夫不等式

$$P\{|X-\mu|\geqslant 3\sigma\}\leqslant\underline{\qquad\qquad}.\qquad(1989\text{ 年四})$$

解　契比雪夫不等式为

$$P\{|X-\mu|\geqslant\varepsilon\}\leqslant\sigma^2/\varepsilon^2,$$

将$\varepsilon=3\sigma$代入即得$P\{|X-\mu|\geqslant 3\sigma\}\leqslant\sigma^2/(9\sigma^2)=1/9$.

3. 某保险公司多年的统计资料表明：在索赔户中被盗索赔户占20%. 以X表示在随意抽查的100个索赔户中因被盗向保险公司索赔的户数.

(1) 写出X的概率分布；

(2) 利用德莫弗-拉普拉斯中心极限定理，求被盗索赔户不少于14户且不多于30户的概率.　　(1988 年四)

解 (1) $X\sim B(100,0.2)$,所以

$$P\{X=k\}=C_{100}^{k}\times 0.2^{k}\times 0.8^{100-k},\quad k=0,1,\cdots,100.$$

(2) $P\{14\leqslant X\leqslant 30\}$

$$=\Phi\left(\frac{30-100\times 0.2}{\sqrt{100\times 0.2\times 0.8}}\right)-\Phi\left(\frac{14-100\times 0.2}{\sqrt{100\times 0.2\times 0.8}}\right)$$

$$=\Phi(2.5)-\Phi(-1.5)=\Phi(2.5)+\Phi(1.5)-1$$

$$\xlongequal{\text{查表}}0.994+0.933-1=0.927.$$

4. 假设 $X_1,X_2,\cdots,X_n$ 是来自总体 X 的简单随机样本:已知 $EX^k=\alpha_k$ $(k=1,2,3,4)$,证明:当 n 充分大时,随机变量

$$Z_n=\frac{1}{n}\sum_{i=1}^{n}X_i^2$$

近似服从正态分布,并指出其分布参数. (1996 年四)

证 依题意 $X_1,X_2,\cdots,X_n$ 相互独立且同分布,可见 $X_1^2,X_2^2,\cdots,X_n^2$ 也相互独立且同分布. 由 $E(X^k)=\alpha_k$ $(k=1,2,3,4)$,有

$$E(X_i^2)=\alpha_2,\quad D(X_i^2)=E(X_i^4)-[E(X_i^2)]^2=\alpha_4-\alpha_2^2,$$

$$E(Z_n)=\frac{1}{n}\sum_{i=1}^{n}E(X_i^2)=\alpha_2,$$

$$D(Z_n)=\frac{1}{n^2}\sum_{i=1}^{n}D(X_i^2)=\frac{\alpha_4-\alpha_2^2}{n}.$$

因此,根据中心极限定理

$$U_n=(Z_n-\alpha_2)\Big/\sqrt{(\alpha_4-\alpha_2^2)/n}\sim N(0,1),$$

即当 n 充分大时,Z_n 近似服从 $N(\alpha_2,(\alpha_4-\alpha_2^2)/n)$.

5. 设随机变量 X 的方差为 2,则根据契比雪夫不等式有估计 $P\{|X-E(X)|\geqslant 2\}\leqslant$______. (2001 年一)

解 由契比雪夫不等式

$$P\{|X-\mu|\geqslant\varepsilon\}\leqslant\sigma^2/\varepsilon^2,$$

将 $\sigma^2=2,\varepsilon=2$ 代入即得 1/2.

6. 设随机变量 X 和 Y 的数学期望分别为 -2 和 2,方差分别为

1 和 4，而相关系数为 -0.5，则根据契比雪夫不等式，有

$$P\{|X+Y|\geqslant 6\}\leqslant\underline{\qquad\qquad}.$$ (2001 年三)

解 $E(X+Y)=E(X)+E(Y)=0$，

$$\begin{aligned}D(X+Y)&=D(X)+D(Y)+2\text{cov}(XY)\\&=D(X)+D(Y)+2\rho_{XY}\sqrt{D(X)}\sqrt{D(Y)}\\&=1+4-2\times 0.5\times 2=3,\end{aligned}$$

所以 $$P\{|X+Y|\geqslant 6\}\leqslant 3/36=1/12.$$

7. 设随机变量 X 和 Y 的数学期望都是2，方差分别为1 和4，而相关系数为 0.5，则根据契比雪夫不等式，有

$$P\{|X-Y|\geqslant 6\}\leqslant\underline{\qquad\qquad}.$$ (2001 年四)

解 $E(X-Y)=E(X)-E(Y)=0$，

$$\begin{aligned}D(X-Y)&=D(X)+D(Y)-2\text{cov}(X,Y)\\&=D(X)+D(Y)-2\rho_{XY}\sqrt{D(X)}\sqrt{D(Y)}\\&=1+4-2\times 0.5\times 2=3,\end{aligned}$$

所以 $$P\{|X-Y|\geqslant 6\}\leqslant 3/36=1/12.$$

8. 一生产线生产的产品成箱包装，每箱的重量是随机的. 假设每箱平均重 50 kg，标准差为 5 kg. 若用最大载重量为 5 t 的汽车承运，试利用中心极限定理说明：每辆车最多可以装多少箱，才能保障不超载的概率大于 0.977 ($\Phi(2)=0.977$，其中 $\Phi(x)$ 是标准正态分布函数). (2001 年三、四)

解 设 X_i $(i=1,2,\cdots,n)$ 是装运的第 i 箱的重量(单位：kg)，n 是所求箱数. 由条件可以把 $X_1,X_2,\cdots,X_n$ 视为相互独立且同分布随机变量，而 n 箱的总重量

$$T_n=X_1+X_2+\cdots+X_n$$

是相互独立且同分布随机变量之和.

由条件知

$$E(X_i)=50\ \sqrt{D(X_i)}=5,\quad E(T_n)=50n\ \sqrt{D(T_n)}=5\sqrt{n}.$$

根据列维-林德伯格中心极限定理，T_n 近似服从正态分布

$N(50n,25n)$. 箱数 n 取决于条件

$$P\{T_n \leqslant 5000\}=P\left\{\frac{T_n-50n}{5\sqrt{n}} \leqslant \frac{5000-50n}{5\sqrt{n}}\right\}$$

$$\approx \Phi\left(\frac{1000-10n}{\sqrt{n}}\right)>0.977=\Phi(2).$$

由此可见
$$\frac{1000-10n}{\sqrt{n}}>2,$$

从而 $n<98.0199$，即最多可装 98 箱.

9. 设随机变量 $X_1,X_2,\cdots,X_n$ 相互独立，$S_n=X_1+X_2+\cdots+X_n$，则根据列维-林德伯格中心极限定理，当 n 充分大时，S_n 近似服从正态分布，只要 $X_1,X_2,\cdots,X_n$(　　).

(A) 有相同的数学期望；　　(B) 有相同的方差；

(C) 服从同一指数分布；　　(D) 服从同一离散型分布.

(2002 年四)

解　选(C). 因为列维-林德伯格中心极限定理要求 $X_1,X_2,\cdots,X_n$ 相互独立且同分布，有有限的数学期望与方差，$E(X_i)=\mu$，所以

$$D(X_i)=\sigma^2\neq 0 \quad (i=1,2,\cdots,n).$$

第六章　数理统计的基本概念

第一节　随 机 样 本

主 要 内 容

1. 总体与个体

研究对象的某项数量指标的全体称为总体.总体中的每个元素称为个体.总体按所含个体的多少分为有限总体与无限总体.总体通常用X来表示,总体的每个个体的取值是随机的,在客观上有一定的分布,所以X是一个随机变量.X的分布函数与数字特征就称为总体的分布函数与数字特征.

2. 简单随机样本

设X是一个具有分布函数F的随机变量,从总体X中随机抽取的有同一分布F的n个相互独立的随机变量$X_1,X_2,\cdots,X_n$,称为总体X的一个容量为n的简单随机样本.按样本实行的一次抽取的具体值,记为$x_1,x_2,\cdots,x_n$,称为一组样本观察值.

$$F^*(x_1,x_2,\cdots,x_n)=\prod_{i=1}^{n}F(x_i)$$

称为样本$X_1,X_2,\cdots,X_n$的联合分布函数.

$$f^*(x_1,x_2,\cdots,x_n)=\prod_{i=1}^{n}f(x_i)\ (若存在)$$

称为样本$X_1,X_2,\cdots,X_n$的联合概率密度.

3. 统计量

设 $X_1, X_2, \cdots, X_n$ 是来自总体 X 的一个样本，$g(X_1,X_2,\cdots,X_n)$是 $X_1,X_2,\cdots,X_n$ 的一个不含任何未知参数的连续函数，称$g(X_1,X_2,\cdots,X_n)$是一个统计量. 统计量也是随机变量.

4. 一些常用统计量

(1) 样本均值 $\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$.

(2) 样本方差

$$S^2=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2=\frac{1}{n-1}\left(\sum_{i=1}^{n}X_i^2-n\overline{X}^2\right).$$

(3) 样本标准差 $S=\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2}$.

(4) 样本k阶(原点)矩 $A_k=\frac{1}{n}\sum_{i=1}^{n}X_i^k,\ k=1,2,\cdots$.

(5) 样本k阶中心矩 $B_k=\frac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})^k,\ k=1,2,\cdots$.

它们的观察值仍分别称样本均值、样本方差、样本标准差、标本k阶矩、样本k阶中心矩.

5. 经验分布函数

若总体为X，$X_1,X_2,\cdots,X_n$是X的一个样本，将样本的一组观察值$x_1,x_2,\cdots,x_n$按大小排成顺序统计量$x_1^*<x_2^*<\cdots<x_n^*$，则

$$F_n(x)=\begin{cases}0, & x<x_1^*,\\ k/n, & x_k^*\leqslant x<x_{k+1}^*,\ k=1,2,\cdots,n-1,\\ 1, & x>x_n^*.\end{cases}$$

称$F_n(x)$为经验分布函数，其图形为一阶跃曲线.

疑 难 解 析

1. 为什么可以把总体看成一个随机变量？

答 当总体表示某项数量指标时，对于X的每个个体，都有

一个对应的取值.这个取值有一定的分布，而且具有随机性，所以 X 是一个随机变量.因此，对总体的研究就转化为对随机变量的研究，了解了随机变量 X，也就了解了总体. X 的分布函数和数字特征就是总体的分布函数和数字特征.

2. 简单随机样本有什么特点？有什么意义？

答 要了解一个总体，当然最好是了解每一个个体，但这样太费时间，代价也太高，因此，用抽取样本的方式来了解是最好的选择.为了使样本 $X_1,X_2,\cdots,X_n$ 具有充分的代表性，应该：(1) $X_1,X_2,\cdots,X_n$ 相互独立；(2) X 中每个个体被抽到的机会相等.满足这两个条件的样本就是简单随机样本.

因为简单随机样本具有充分的代表性，所以用从简单随机样本得到的信息去推断总体有可靠的依据.而样本分布函数与样本概率密度又为引进统计量，将样本中信息集中起来，进而为推断总体提供了有效的途径.

3. 统计量有什么意义？为什么统计量中不能含有未知参数？

答 样本是总体的反映，又是进行统计推断的依据.但样本反映的信息是零乱的、无序的和分散的，所以要针对不同的问题构造样本的不同函数，将信息集中起来，以便进行统计推断和研究分析，使之更易揭示问题的本质.统计量就是样本的不含未知参数的连续函数.例如，$\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$ 是一个统计量，不含任何未知参数，它排除了关于 σ^2 的信息，集中了关于 μ 的信息.

连续性是为了保证统计量仍然是随机变量而提出的，同时也为以后用数学分析方法研究统计量(如极大似然函数)提供了条件.

不含未知参数，则统计量只与样本有关，而与总体无关.若含有未知参数，则无法依靠样本观察值来求未知参数的估计值，因而失去利用统计量估计未知参数的作用，这是违背我们引进统计量的初衷的.

4. 经验分布函数与分布函数有什么关系?

答 经验分布函数是由总体 X 的一个样本 $X_1,X_2,\cdots,X_n$ 的一次实现 $x_1,x_2,\cdots,x_n$ 构造的一个函数. 它既是 X 的函数,又是顺序统计量 $X_1^*,X_2^*,\cdots,X_n^*$ 的函数. 显然,它对不同的样本,不同次的实现是不唯一的.

经验分布函数 $F_n(x)$ 在样本的一组观察值确定后就确定了. 它具有:(1) 单调不减性. 当 $x_1<x_2$ 时,$F_n(x_1)<F_n(x_2)$. (2) 有界性. $0\leqslant F_n(x)\leqslant 1$. (3) 右连续性. $F(x+0)=F(x)$. 因此,经验分布函数 $F_n(x)$ 是随机变量的分布函数. 事实上,它可以视为一个概率分布为

$$P\{X=x_k\}=1/n,\quad k=1,2,\cdots,n$$

的离散型随机变量的分布函数.

经验分布函数 $F_n(x)$ 的值依赖于样本观察值,不含未知参数,是一个统计量. 又因为对每一组样本观察值,有不同的 $F_n(x)$,所以经验分布函数又是一个随机变量. 由格里文科(W. Glivenko)定理,当 $n\to\infty$ 时,$F_n(x)$ 关于 x 依概率收敛于 $F(x)$. 因此,当 n 充分大($n\geqslant 50$,最好 $n\geqslant 100$)时,用 $F_n(x)$ 代替 $F(x)$ 是可行的.

方法、技巧与典型例题分析

本节的例题主要是加深对数理统计基本概念的理解,利用概率论中学习过的知识和数理统计基本概念的定义进行计算,或者验证一些命题. 一般可以直接计算.

一、总体、样本及其分布、样本的数字特征

例 1 某厂生产的某种电器的使用寿命服从指数分布,参数 λ 未知. 为此,抽查了 n 件电器,测量其使用寿命,试确定本问题的总体、样本及样本的分布.

解 总体是这种电器的使用寿命,其概率密度为

$$f(x)=\begin{cases}\lambda e^{-\lambda x}, & x>0,\\ 0, & x\leqslant 0\end{cases}(\lambda\text{未知}).$$

样本 $X_1,X_2,\cdots,X_n$ 是 n 件某种电器的使用寿命，抽到的 n 件电器的使用寿命是样本的一组观察值. 样本 $X_1,X_2,\cdots,X_n$ 相互独立，来自同一总体 X，所以样本的联合密度为

$$f(x_1,x_2,\cdots,x_n)=\begin{cases}\lambda^n e^{-\lambda(x_1+x_2+\cdots+x_n)}, & x_1,x_2,\cdots,x_n>0,\\ 0, & \text{其它}.\end{cases}$$

例 2 设 $x_1,x_2,\cdots,x_n$ 是总体 X 的一组样本观察值，则使 $\sum\limits_{i=1}^{n}(x_i-a)^2$ 取最小值的 a 等于什么？

解 设 $f(a)=\sum\limits_{i=1}^{n}(x_i-a)^2$，求其极值. 求导，即

$$f'_a(a)=-2\sum_{i=1}^{n}(x_i-a),$$

令上式等于零，得 $$a=\frac{1}{n}\sum_{i=1}^{n}x_i.$$

又 $$f''_a(a)=2>0,$$

故当 $a=\frac{1}{n}\sum\limits_{i=1}^{n}x_i$ 时，$\sum\limits_{i=1}^{n}(x_i-a)^2$ 取最小值.

例 3 设随机变量 $X\sim N(1,4)$，$X_1,X_2,\cdots,X_{100}$ 是 X 的一个样本，$\overline{X}=\frac{1}{100}\sum\limits_{i=1}^{100}X_i$. 若 $Y=a\overline{X}+b\sim N(0,1)$，求 a,b 的值.

解 $X\sim N(1,4)$，则 $\overline{X}\sim N(1,1/25)$. 而

$$Y=a\overline{X}+b\sim N(a+b,a^2/25),$$

又 $Y\sim N(0,1)$，所以，$a=\pm 5$，$b=\mp 5$.

例 4 设 $X_1,X_2,\cdots,X_n$ 为 X 的一个样本，$X\sim N(\mu,\sigma^2)$，求 $\overline{X}$ 的分布，$\sum\limits_{i=1}^{n}a_iX_i$ ($a_i\neq 0$，常数)的分布.

解 $X\sim N(\mu,\sigma^2)$，则 $E(X)=\mu$，$D(X)=\sigma^2$，故

$$E(\overline{X})=\frac{1}{n}\sum_{i=1}^{n}E(X_i)=\frac{1}{n}n\mu=\mu,$$

$$D(\overline{X})=\frac{1}{n^2}\sum_{i=1}^{n}D(X_i)=\frac{1}{n^2}n\sigma^2=\frac{1}{n}\sigma^2,$$

即

$$\overline{X}\sim N(\mu,\sigma^2/n).$$

$$E\Big[\sum_{i=1}^{n}a_iX_i\Big]=\sum_{i=1}^{n}a_iE(X_i)=\mu\sum_{i=1}^{n}a_i,$$

$$D\Big[\sum_{i=1}^{n}a_iX_i\Big]=\sum_{i=1}^{n}a_i^2D(X_i)=\sigma^2\sum_{i=1}^{n}a_i^2,$$

即

$$\sum_{i=1}^{n}a_iX_i\sim N\Big(\mu\sum_{i=1}^{n}a_i,\sigma^2\sum_{i=1}^{n}a_i^2\Big).$$

例5 总体 X 的一组容量为5的样本观察值为8,2,5,3,7,求样本均值 $\overline{x}$、样本方差 S^2、样本二阶中心矩 b_2 及经验分布函数 $F_5(x)$.

解 由定义 $\overline{x}=\frac{1}{5}(8+2+5+3+7)=5$,

$$S^2=\frac{1}{4}[3^2+(-3)^2+0^2+(-2)^2+2^2]=6.5,$$

$$b_2=\frac{1}{5}[3^2+(-3)^2+0^2+(-2)^2+2^2]=5.2.$$

$$F_5(x)=\begin{cases}0, & x<2,\\ 0.2, & 2\leqslant x<3,\\ 0.4, & 3\leqslant x<5,\\ 0.6, & 5\leqslant x<7,\\ 0.8, & 7\leqslant x<8,\\ 1, & x\geqslant 8.\end{cases}$$

例6 设总体 X 的分布函数为 $F(x)$,经验分布函数为 $F_n(x)$,证明:

$$E[F_n(x)]=F(x),\quad D[F_n(x)]=\frac{1}{n}F(x)[1-F(x)].$$

证 引入随机变量 $nF_n(x)$,则

$$P\{nF_n(x)=k\}=C_n^k F^k(x)[1-F(x)]^{n-k},\ k=0,1,\cdots,n,$$

其中$F(x)$是X的分布函数,所以,由二项分布$B(n,p)$的数学期望与方差公式知

$$E[nF_n(x)]=nF(x)\Longrightarrow E[F_n(x)]=F(x),$$

$$D[nF_n(x)]=nF(x)[1-F(x)]$$

$$\Longrightarrow D[F_n(x)]=\frac{1}{n}F(x)[1-F(x)].$$

例7 设总体$X\sim \mathrm{e}(\lambda)$,$X_1,X_2,\cdots,X_n$是$X$的一个样本,求$E(\overline{X})$,$E(S^2)$.

解 $$E(X)=1/\lambda,\quad D(X)=1/\lambda^2,$$

$$E(X^2)=D(X)+[E(X)]^2=2/\lambda^2.$$

由于$X_1,X_2,\cdots,X_n$相互独立且同分布,故

$$E(\overline{X})=\frac{1}{n}\sum_{i=1}^{n}E(X_i)=\frac{1}{n}\cdot\frac{n}{\lambda}=\frac{1}{\lambda},$$

$$E(\overline{X}^2)=D(\overline{X}^2)+[E(\overline{X})]^2=\frac{1}{n^2}\cdot\frac{n}{\lambda^2}+\frac{1}{\lambda^2}=\frac{n+1}{n\lambda^2},$$

$$E(S^2)=E\left[\frac{1}{n-1}\left(\sum_{i=1}^{n}X_i^2-n\overline{X}^2\right)\right]$$

$$=\frac{1}{n-1}\left[\sum_{i=1}^{n}E(X_i^2)-nE(\overline{X}^2)\right]$$

$$=\frac{1}{n-1}\left(n\,\frac{2}{\lambda^2}-n\,\frac{n+1}{n\lambda^2}\right)=\frac{1}{\lambda^2}.$$

例8 设$\overline{X}_n$和S_n^2分别是样本$X_1,X_2,\cdots,X_n$的样本均值及样本方差.若添加一次试验,则样本扩展为$X_1,X_2,\cdots,X_n,X_{n+1}$,其样本均值和样本方差分别$\overline{X}_{n+1}$和$S_{n+1}^2$.证明下列递推公式成立:

$$\overline{X}_{n+1}=\overline{X}_n+\frac{1}{n+1}(X_{n+1}-\overline{X}_n),$$

$$S_{n+1}^2=\frac{n-1}{n}S_n^2+\frac{1}{n+1}(X_{n+1}-\overline{X}_n)^2.$$

证 用定义证.

$$\overline{X}_{n+1}=\frac{1}{n+1}[(X_1+X_2+\cdots+X_n)+X_{n+1}]$$
$$=\frac{n}{n+1}\cdot\frac{1}{n}(X_1+X_2+\cdots+X_n)+\frac{1}{n+1}X_{n+1}$$
$$=\frac{n}{n+1}\overline{X}_n+\frac{1}{n+1}X_{n+1}=\overline{X}_n+\frac{1}{n+1}(X_{n+1}-\overline{X}_n).$$
$$S_{n+1}^2=\frac{1}{n}\sum_{i=1}^{n+1}(X_i-\overline{X}_{n+1})^2$$
$$=\frac{1}{n}\sum_{i=1}^{n+1}\left(X_i-\frac{n}{n+1}\overline{X}_n-\frac{1}{n+1}X_{n+1}\right)^2$$
$$=\frac{1}{n}\sum_{i=1}^{n}\left[(X_i-\overline{X}_n)+\left(\frac{1}{n+1}\overline{X}_n-\frac{1}{n+1}X_{n+1}\right)\right]^2$$
$$+\frac{1}{n}\left(X_{n+1}-\frac{n}{n+1}\overline{X}_n-\frac{1}{n+1}X_{n+1}\right)^2$$
$$=\frac{n-1}{n}\cdot\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X}_n)^2+2(\overline{X}_n-X_{n+1})$$
$$\cdot\sum_{i=1}^{n}\frac{X_i-\overline{X}_n}{n(n+1)}+\frac{(\overline{X}_n-X_{n+1})^2}{(n+1)^2}+\frac{n}{n+1}(\overline{X}_n-X_{n+1})^2$$
$$=\frac{n-1}{n}S_n^2+\frac{1}{n+1}(X_{n+1}-\overline{X}_n)^2.$$

以上公式阐明，当样本容量再增加一个时，可以利用前 n 个数据得出的均值和方差添加新的数据得到新的样本均值与方差.

例 9 设有 N 个产品，其中有 M 个次品，进行放回抽样. 定义 X_i 如下：

$$X_i=\begin{cases}1, & \text{第 } i \text{ 次取得次品},\\ 0, & \text{第 } i \text{ 次取得正品}.\end{cases}$$

求样本 $X_1,X_2,\cdots,X_n$ 的联合分布.

解 因为是放回抽样，所以 $X_1,X_2,\cdots,X_n$ 相互独立且同分布，且

$$P\{X_i=1\}=\frac{M}{N},\quad P\{X_i=0\}=1-\frac{M}{N},$$

因此 $X_1,X_2,\cdots,X_n$ 的联合分布为

$$P\{X_1=x_1,\cdots,X_n=x_n\}=(M/N)^{\sum\limits_{i=1}^{n}x_i}(1-M/N)^{n-\sum\limits_{i=1}^{n}x_i}.$$

例 10 设$\overline{X}$ 和S_x^2 是样本$X_1,X_2,\cdots,X_n$ 的样本均值和样本方差，作数据变换$y_i=(x_i-a)/c,i=1,2,\cdots,n$. 设$\overline{Y}$ 和S_Y^2 为样本Y_1，$Y_2,\cdots,Y_n$ 的样本均值和样本方差，证明：

(1) $\overline{X}=a+c\overline{Y}$；　　(2) $S_X^2=c^2S_Y^2$.

证 (1)
$$\overline{Y}=\frac{1}{n}\sum_{i=1}^{n}Y_i=\frac{1}{n}\sum_{i=1}^{n}\frac{X-a}{c}$$
$$=\frac{1}{n}\sum_{i=1}^{n}\frac{X_i}{c}-\frac{a}{c}=\frac{\overline{X}}{c}-\frac{a}{c},$$

即
$$\overline{X}=a+c\overline{Y}.$$

(2)
$$S_x^2=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2=\frac{1}{n-1}\sum_{i=1}^{n}[(Y_i+a)-(a+c\overline{Y})]^2$$
$$=\frac{1}{n-1}\sum_{i=1}^{n}c^2(Y_i-\overline{Y})^2=c^2S_Y^2.$$

这两个式子可以用来简化运算，特别是数据数值较大或者带有小数时，应用这两个式子可以给计算带来很大方便.

例 11 设随机变量X 的概率密度为$f(x)$，数学期望$E(X)=\mu,D(X)=\sigma^2$ 存在，$X_1,X_2,\cdots,X_n$ 是X 的一个样本，$\overline{X}$ 是样本均值，则有(　　)成立.

(A) $\overline{X}\sim f(x)$；　　(B) $\min\limits_{1\leqslant i\leqslant n}\{X_i\}-f(x)$；

(C) $\max\limits_{1\leqslant i\leqslant n}\{X_i\}\sim f(x)$；　　(D) $(X_1,X_2,\cdots,X_n)\sim\prod\limits_{i=1}^{n}f(x_i)$.

解 选(D). 由于$X_1,X_2,\cdots,X_n$ 相互独立且同分布，因而联合密度等于各边缘密度之乘积.

因为$X\sim f(x)$，所以$D(\overline{X})=\frac{1}{n}D(X)=\frac{1}{n}\sigma^2$，即所以(A)不成立.

令$Y=\min\limits_{1\leqslant i\leqslant n}\{X_i\}$，则

$F_Y(x)=P\{Y\leqslant x\}=1-P\{Y\geqslant x\}=1-P\{\min\{X_i\}\geqslant x\}$

$$=1-P\left\{\prod_{i=1}^{n}\{X_i\geqslant x\}\right\}=1-[1-F_X(x)]^n\neq F_X(x),$$

所以(B)不成立.

令 $Z=\max\{X_i\}$，则

$$F_Z(x)=P\{Z\leqslant x\}=P\{\max_{1\leqslant i\leqslant n}\{X_i\}\leqslant x\}=P\left\{\prod_{i=1}^{n}\{X_i\leqslant x\}\right\}$$

$$=\prod_{i=1}^{n}P\{X_i\leqslant n\}=[F_X(x)]^n\neq F_X(x),$$

所以(C)不成立.

二、样本统计量的概率与样本容量的确定

当总体已知时，对于从总体中抽取的样本，往往要计算样本统计量落入某区间内的概率，或是在概率已知的条件下计算样本容量要取多大. 其基本解法是：由总体的分布确定样本统计量的分布，然后查表，求出概率；或者由概率与样本容量的关系式确定 n.

例12 设总体 $X\sim N(20,3)$，从 X 中抽取两个样本 $X_1,X_2,\cdots,X_{10}$ 和 $Y_1,Y_2,\cdots,Y_{15}$，求概率 $P\{|\overline{X}-\overline{Y}|>3\}$.

解 因为 $X_1,X_2,\cdots,X_{10}$ 和 $Y_1,Y_2,\cdots,Y_{15}$ 相互独立且同分布，所以 $\overline{X}\sim N(20,3/10)$，$\overline{Y}\sim N(20,0.2)$，于是 $\overline{X}-\overline{Y}\sim N(0,0.5)$.

$$P\{|\overline{X}-\overline{Y}|>3\}=P\{|\overline{X}-\overline{Y}|/\sqrt{0.5}>3/\sqrt{0.5}\}$$
$$=1-P\{|\overline{X}-\overline{Y}|/\sqrt{0.5}<3/\sqrt{0.5}\}$$
$$=2[1-\Phi(3/\sqrt{0.5})]=2(1-0.6628)$$
$$=0.6744\ (\text{查正态分布表}).$$

例13 设 $X_1,X_2,\cdots,X_n$ 为总体 $X\sim B(1,p)$ 的一个样本，p 未知，则 $P\{\overline{X}=k/n\}=(\quad)$.

(A) p;　　(B) $1-p$;

(C) $C_n^k p^k(1-p)^{n-k}$;　　(D) $C_n^k(1-p)^k p^{n-k}$.

解 选(C). 因为 $X_1,X_2,\cdots,X_n$ 相互独立且同分布，所以 $\sum_{i=1}^{n}X_i\sim B(n,p)$，于是

$$P\left\{\sum_{i=1}^{n}X_i=k\right\}=C_n^k p^k(1-p)^{n-k},$$

而
$$P\left\{\overline{X}=\frac{k}{n}\right\}=P\left\{\sum_{i=1}^{n}X_i=k\right\}.$$

例 14 设总体 $X\sim N(\mu,\sigma^2)$，若要以 99.7%的概率保证偏差 $|\overline{X}-\mu|<0.1$，问：在 $\sigma^2=0.5$ 时，样本容量 n 应取多大？

解 要使 $|\overline{X}-\mu|<0.1$，即要

$$P\{|\overline{X}-\mu|<0.1\}=P\{|\overline{X}-\mu|/\sqrt{0.5/n}<0.1/\sqrt{0.5/n}\}$$
$$=2\Phi(0.141\sqrt{n})-1=0.997.$$

查正态分布表知 $0.141\sqrt{n}=2.97$，所以 $n=444$.

例 15 设总体 $X\sim N(3.4,6^2)$，$X_1,X_2,\cdots,X_n$ 为 X 的一个简单随机样本，要使 $P\{1.4<\overline{X}<5.4\}\geqslant 0.95$，样本容量 n 应取多大？

解 由题设知，$(\overline{X}-3.4)/(6/\sqrt{n})\sim N(0,1)$，故

$$P\{1.4<\overline{X}<5.4\}=P\{|\overline{X}-3.4|<2\}$$
$$=P\left\{\left|\frac{\overline{X}-3.4}{6}\right|\sqrt{n}<\frac{\sqrt{n}}{3}\right\}$$
$$=2\Phi\left(\frac{\sqrt{n}}{3}\right)-1\geqslant 0.95$$
$$\Longrightarrow\Phi\left(\frac{\sqrt{n}}{3}\right)\geqslant 0.975.$$

查正态分布表知 $\sqrt{n}/3\geqslant 1.96$，所以 $n=35$.

例 16 设总体 $X\sim B(1,p)$，$X_1,X_2,\cdots,X_n$ 为 X 的一个样本，p 未知，问：对每个 p $(0<p<1)$，n 应取多大，才能保证

$$E[(\overline{X}-p)^2]\leqslant 0.01?$$

解 因为 $E(X)=p$，$D(X)=pq$，$E(\overline{X})=p$，$D(\overline{X})=pq/n$，所以，要使

$$E[(\overline{X}-p)]^2=D(\overline{X})=pq/n\leqslant 0.01,$$

应有 $n\geqslant 100pq=100(1-p)p$. 但由不等式性质知，$p(1-p)\leqslant 1/4$，故 $n\geqslant 25$.

例 17 设 $X_1,X_2,\cdots,X_n$ 是来自总体 $X\sim N(\mu,\sigma^2)$ 的一个样本,求满足下式的最小 n 值

$$P\{|S^2-\sigma^2|\leqslant\sigma^2/2\}\geqslant 0.8.$$

解 将不等式变形,得

$$P\left\{|S^2-\sigma^2|\leqslant\frac{\sigma^2}{2}\right\}=P\left\{\left|\frac{S^2}{\sigma^2}-1\right|\leqslant\frac{1}{2}\right\}$$

$$=P\left\{\left|\frac{S^2}{\sigma^2}-1\right|\Big/\sqrt{\frac{2}{n-1}}<\frac{1}{2}\frac{\sqrt{n-1}}{\sqrt{2}}\right\}$$

$$=2\Phi\left(\frac{\sqrt{n-1}}{2\sqrt{2}}\right)-1\geqslant 0.8$$

$$\Longrightarrow\Phi\left(\frac{\sqrt{n-1}}{2\sqrt{2}}\right)\geqslant 0.9.$$

查正态分布表知 $\dfrac{\sqrt{n-1}}{2\sqrt{2}}\geqslant 1.28$,故 $n=14$.

例 18 设总体 $X\sim N(\mu,4)$,样本 $X_1,X_2,\cdots,X_n$ 来自 X,样本容量 n 取多大时,有

(1) $E(|\overline{X}-\mu|^2)\leqslant 0.1$; (2) $P\{|\overline{X}-\mu|\leqslant 0.1\}\geqslant 0.95$?

解 (1) 因为 $E(|\overline{X}-\mu|^2)=D(\overline{X})=D(X)/n$,所以

$$E(|\overline{X}-\mu|^2)\leqslant 0.1\Longrightarrow 4/n\leqslant 0.1\Longrightarrow n\geqslant 40,$$

故当样本容量 $n=40$ 时,$E(|\overline{X}-\mu|^2)\leqslant 0.1$.

(2) 由中心极限定理,要使

$$P\{|\overline{X}-\mu|<0.1\}$$

$$=P\left\{|\overline{X}-\mu|\Big/\sqrt{D(\overline{X})}<0.1\Big/\sqrt{D(\overline{X})}\right\}\geqslant 0.95,$$

必须

$$2\Phi\left(0.1\Big/\sqrt{D(\overline{X})}\right)-1\geqslant 0.95\Longrightarrow\Phi\left(0.1\Big/\sqrt{D(\overline{X})}\right)\geqslant 0.975.$$

因为 $$D(\overline{X})=D(X)/n=4/n,$$

所以,查正态分布表知 $0.05\sqrt{n}\geqslant 1.96$,即 $n=1537$.

例 19 设总体 X 的密度函数为

$$f(x)=\begin{cases}|x|, & |x|<1,\\ 0, & \text{其它},\end{cases}$$

$X_1,X_2,\cdots,X_{50}$为取自X的一个样本，求：

(1) $E(\overline{X})$和$D(\overline{X})$； (2) $E(S^2)$； (3) $P\{|\overline{X}|>0.02\}$.

解 $E(X)=\int_{-1}^{1}|x|x\mathrm{d}x=0$,

$$D(X)=E(X^2)=\int_{-1}^{1}|x|x^2\mathrm{d}x=2\int_0^1 x^3\mathrm{d}x=\frac{1}{2}.$$

(1) $E(\overline{X})=E(X)=0$, $D(\overline{X})=\frac{1}{n}D(X)=\frac{1}{2n}=\frac{1}{100}$.

(2) $$E(S^2)=E\left[\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2\right]=\frac{1}{n-1}E\left[\sum_{i=1}^{n}(X_i-\overline{X})^2\right]$$

$$=\frac{1}{n-1}D(\overline{X})=\frac{1}{n-1}\cdot\frac{1}{2n}=\frac{1}{2n(n-1)},$$

故 $$E(S^2)=\frac{1}{2\times 50\times 49}=\frac{1}{4900}.$$

(3) $$P\{|\overline{X}|>0.02\}=1-P\{|\overline{X}|<0.002\}$$

$$=1-P\left\{\left|\frac{\overline{X}-\mu}{\sqrt{D(x)}}\right|\leqslant\frac{0.02-\mu}{\sqrt{D(X)}}\right\}$$

$$=1-P\{10|\overline{X}|\leqslant 0.2\}$$

$$=2-2\Phi(0.2)$$

$$=0.8414.$$

第二节 正态总体下的抽样分布

主要内容

1. 样本均值$\overline{X}$的分布

设$X\sim N(\mu,\sigma^2)$，$X_1,X_2,\cdots,X_n$是X的一个样本，则样本均值

$\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$ 是统计量，且

$$\overline{X}\sim N(\mu,\sigma^2/n) \quad 或 \quad (\overline{X}-\mu)/\sqrt{\sigma^2/n}\sim N(0,1).$$

$(\overline{X}-\mu)/\sqrt{\sigma^2/n}$ 也称为 U 统计量.

2. χ^2 分布

设 $X\sim N(0,1)$，$X_1,X_2,\cdots,X_n$ 是 X 的一个样本，则

$$\chi^2=X_1^2+X_2^2+\cdots+X_n^2$$

的分布称为服从自由度为 n 的 χ^2 分布，记为 $\chi^2\sim\chi^2(n)$. $\chi^2(n)$ 分布的概率密度函数为

$$f(y)=\begin{cases}\frac{1}{2^{n/2}\Gamma(n/2)}y^{n/2-1}\mathrm{e}^{-y/2}, & y>0,\\ 0 & 其它.\end{cases}$$

若 $X\sim N(\mu,\sigma^2)$，$X_1,X_2,\cdots,X_n$ 是 X 的一个样本，则

$$\chi^2=\frac{1}{\sigma^2}\sum_{i=1}^{n}(X_i-\mu)^2\sim\chi^2(n).$$

χ^2 分布具有以下性质：

(1) 若 $\chi^2\sim\chi^2(n)$，则 $E(\chi^2)=n$，$D(\chi^2)=2n$，

(2) 若 $X_1=\chi_1^2\sim\chi^2(n_1)$，$X_2=\chi_2^2\sim\chi^2(n_2)$，且 X_1,X_2 相互独立，则

$$X_1+X_2=\chi_1^2+\chi_2^2\sim\chi^2(n_1+n_2).$$

此结果可以推广到有限个 χ^2 分布相加的情形.

3. t 分布

设总体 $X\sim N(0,1)$，$Y\sim\chi^2(n)$，且 X,Y 相互独立，则称随机变量 $t=\frac{X}{\sqrt{Y/n}}$ 服从自由度为 n 的 t 分布，记为 $t\sim t(n)$. $t(n)$ 分布的概率密度函数为

$$f(t)=\frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\,\Gamma\left(\frac{n}{2}\right)}\left(1+\frac{t^2}{n}\right)^{-\frac{n+1}{2}}, \quad -\infty<t<+\infty.$$

t 分布具有以下性质：

$$\lim_{n\to\infty} f(t)=\frac{1}{\sqrt{2\pi}}e^{-t^2/2}.$$

对 $t(n)$ 的上 α 分位点，由 $f(t)$ 的图形对称性得

$$t_{1-\alpha}(n)=-t_\alpha(n).$$

4. F 分布

设总体 $U\sim\chi^2(n_1)$，$V\sim\chi^2(n_2)$，U 与 V 相互独立，则称统计量 $F=\frac{U/n_1}{V/n_2}$ 服从第一自由度为 n_1，第二自由度为 n_2 的 F 分布，记为 $F\sim F(n_1,n_2)$. $F(n_1,n_2)$ 的概率密度为

$$f(y)=\begin{cases}\dfrac{\Gamma\left(\dfrac{n_1+n_2}{2}\right)}{\Gamma\left(\dfrac{n_1}{2}\right)\Gamma\left(\dfrac{n_2}{2}\right)}\left(\dfrac{n_1}{n_2}\right)^{\frac{n_1}{2}}(y)^{\frac{n_1}{2}-1}\left(1+\dfrac{n_1}{n_2}y\right)^{-\frac{n_1+n_2}{2}}, & y>0,\\ 0, & y\leqslant 0.\end{cases}$$

若 $F\sim F(n_1,n_2)$，则 $1/F\sim F(n_2,n_1)$.

5. 正态总体样本方差 S^2 的分布

设总体 $X\sim N(\mu,\sigma^2)$，$X_1,X_2,\cdots,X_n$ 是 X 的一个样本，$S^2=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2$ 是样本方差，则

(1) $\frac{(n-1)S^2}{\sigma^2}\sim\chi^2(n-1)$；

(2) $\overline{X}$ 与 S^2 相互独立.

由此可以推出：

若 $X_1,X_2,\cdots,X_n$ 是 $X\sim N(\mu,\sigma^2)$ 的一个样本，则

$$\frac{\overline{X}-\mu}{S/\sqrt{n}}\sim t(n-1).$$

若 $X_1,X_2,\cdots,X_{n_1}$ 是 $X\sim N(\mu_1,\sigma^2)$ 的一个样本，$Y_1,Y_2,\cdots,Y_{n_2}$ 是 $Y\sim N(\mu_2,\sigma^2)$ 的一个样本，X,Y 相互独立. $\overline{X},\overline{Y}$ 是两样本的样本均值，S_1^2,S_2^2 是两样本的样本方差，则

$$\frac{(\overline{X}-\overline{Y})-(\mu_1-\mu_2)}{S_W\sqrt{1/n_1+1/n_2}}\sim t(n_1+n_2-2),$$

其中

$$S_W=[(n_1-1)S_1^2+(n_2-1)S_2^2]/(n_1+n_2-2).$$

设$X_1,X_2,\cdots,X_{n_1}$是$X\sim N(\mu_1,\sigma_1^2)$的一个样本，$Y_1,Y_2,\cdots,Y_{n_2}$是$Y\sim N(\mu_2,\sigma_2^2)$的一个样本，它们相互独立，若$S_1^2,S_2^2$是它们的样本方差，则

$$\frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2}\sim F(n_1-1,n_2-1).$$

疑 难 解 析

1. 什么是自由度？怎样计算自由度？

答　自由度通常是指不受任何约束、可以自由变动的变量的个数. 在数理统计概念中，自由度是对随机变量的二次型而言的，因为一个含有n个变量的二次型

$$\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}X_iX_j\quad(a_{ij}=a_{ji};i,j=1,2,\cdots,n)$$

的秩是指对称矩阵$\boldsymbol{A}=(a_{ij})_{n\times n}$的秩，它的大小反映$n$个变量中能自由变动的无约束变量的多少. 所谓自由度，就是二次型的秩.

计算自由度有两种方法，举例如下：求统计量$\sum\limits_{i=1}^{n}(X_i-\overline{X})^2$的自由度.

一种方法是：因为

$$\begin{aligned}\sum_{i=1}^{n}(X_i-\overline{X})^2&=\sum_{i=1}^{n}X_i^2-n\overline{X}^2=\sum_{i=1}^{n}X_i^2-\frac{1}{n}\Big(\sum_{i=1}^{n}X_i\Big)^2\\&=\sum_{i=1}^{n}\Big(1-\frac{1}{n}\Big)X_i^2+\sum_{\substack{i\neq j\\i,j=1}}^{n}\Big(-\frac{1}{n}\Big)X_iX_j\\&=\boldsymbol{X}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{X},\end{aligned}$$

其中 $\boldsymbol{X}=\begin{pmatrix}X_1\\X_2\\\vdots\\X_n\end{pmatrix}$，$\boldsymbol{A}=\begin{pmatrix}1-\frac{1}{n} & -\frac{1}{n} & \cdots & -\frac{1}{n}\\ -\frac{1}{n} & 1-\frac{1}{n} & \ddots & \vdots\\ \vdots & \ddots & \ddots & -\frac{1}{n}\\ -\frac{1}{n} & \cdots & -\frac{1}{n} & 1-\frac{1}{n}\end{pmatrix}$.

通过矩阵的初等变换可以求得 A 的秩为 $n-1$，所以统计量 $\sum\limits_{i=1}^{n}(X_i-\overline{X})^2$ 的自由度为 $n-1$.

另一种简单的说法是，因为统计量的样本容量为 n，统计量中含有 $\overline{X}$，而 $\overline{X}=\frac{1}{n}(X_1+X_2+\cdots+X_n)$ 是一个约束条件，所以统计量的自由度为 $n-1$.

在一般的问题中，通常都采用这种简单的说法. 如 $\frac{(n-1)S^2}{\sigma^2}$，$\frac{\overline{X}-\mu}{S/\sqrt{n}}$ 中因为含 $\overline{X}$ 和 S^2，是一个约束条件，所以自由度为 $n-1$；而 $\frac{\overline{X}-\overline{Y}-(\mu_1-\mu_2)}{S_W\sqrt{1/n_1+1/n_2}}$ 中含有 $\overline{X}$ 和 $\overline{Y}$，有两个约束条件，所以自由度为 n_1+n_2-2.

2. U 分布、t 分布、χ^2 分布和 F 分布等统计量之间有什么联系与区别？

答 这些分布都是正态总体下的抽样分布，都是在正态总体的前提下，用不同的方式构造出来的. 因为构造的形式不同，所得的分布就不同，所以它们既有联系又有区别.

例如，t 分布与标准正态分布十分相似，当 $n\to\infty$ 时，两者没有大的区别；但当 n 较小时，区别就较明显了. 如 t 分布在 $|x|\to\infty$，密度函数是 $|x|^{-n+1}$ 数量级的，而标准正态分布的密度函数是 $e^{-x^2/2}$ 数量级的. 因此，t 分布只有最高到 $(n-1)$ 阶（整数阶）的矩，而标

准正态分布有任意阶矩，且 t 分布的方差（若存在）也比标准正态分布的方差大.

3. 什么是大样本与小样本？

答 在样本容量固定的条件下进行的统计推断和分析问题称为小样本问题.因为样本容量固定时，如能得到有关统计量或样本函数的精确分布，就能较精确地和较满意地讨论和分析各种统计问题.

在样本容量趋于无穷的条件下进行的统计推断和分析问题称为大样本问题.此时能求出有关统计量或样本函数的极限分布，也可以利用极限分布作为近似分布来作统计推断.

所以，大样本与小样本不单纯是以样本容量的大小来区分的，主要是以得到统计量或样本函数的方式（固定容量或极限形式）来区分的.

方法、技巧与典型例题分析

抽样分布问题包括：(1) 判别总体 X 的一个样本的统计量的分布，又含确定分布类型和确定分布中的参数；(2) 计算抽样分布的概率；(3) 在抽样分布的概率已知的情况下，确定抽样的样本容量；(4) 证明抽样分布的等式或不等式.要解决抽样分布问题，必须对 U 分布、χ^2 分布、t 分布和 F 分布的构造十分熟悉，能根据问题条件确定抽样分布的形式、自由度，然后解决其余的问题.

例 1 设 X_1,X_2,X_3,X_4 是来自总体 $X\sim N(0,4)$ 的一个样本，问：a,b 取何值时，$Y=a(X_1-2X_2)^2+b(3X_3-4X_4)^2\sim\chi^2(n)$？并确定 n 的值.

解 因为 X_1,X_2,X_3,X_4 相互独立且同分布，所以

$$E(X_1-2X_2)=0,\quad E(3X_3-4X_4)=0,$$

$$D(X_1-2X_2)=D(X_1)+4D(X_2)=20,$$

$$D(3X_3-4X_4)=9D(X_3)+16D(X_4)=100,$$

于是 $\frac{X_1-2X_2}{\sqrt{20}}\sim N(0,1)$, $\frac{3X_3-4X_4}{10}\sim N(0,1)$,

而且 X_1-2X_2 与 $3X_3-4X_4$ 相互独立. 此时

$$\frac{(X_1-2X_2)^2}{20}+\frac{(3X_3-4X_4)^2}{100}\sim\chi^2(2),$$

从而知 $a=1/20, b=1/100, n=2$.

例2 设总体 $X\sim N(0,1)$, $X_1,X_2,\cdots,X_6$ 为 X 的一个样本,令 $Y=(X_1+X_2+X_3)^2+(X_4+X_5+X_6)^2$,求常数 C,使 CY 服从 χ^2 分布.

解 因为各 X_i 相互独立且同分布,所以

$$X_1+X_2+X_3\sim N(0,3),\quad X_4+X_5+X_6\sim N(0,3),$$

$$(X_1+X_2+X_3)/\sqrt{3}\sim N(0,1),$$

$$(X_4+X_5+X_6)/\sqrt{3}\sim N(0,1).$$

于是

$$\begin{aligned}&(X_1+X_2+X_3)^2+(X_4+X_5+X_6)^2\\&=3\left[\left(\frac{X_1+X_2+X_3}{\sqrt{3}}\right)^2+\left(\frac{X_4+X_5+X_6}{\sqrt{3}}\right)^2\right]\\&=Y_1^2+Y_2^2,\end{aligned}$$

其中

$$Y_1^2/3\sim\chi^2(1),\quad Y_2^2/3\sim\chi^2(1).$$

所以

$$\begin{aligned}\frac{1}{3}(Y_1^2+Y_2^2)&=\frac{1}{3}[(X_1+X_2+X_3)^2+(X_4+X_5+X_6)^2]\\&=\frac{1}{3}Y\sim\chi^2(2),\end{aligned}$$

即

$$C=1/3.$$

例3 设 $X_1,X_2,\cdots,X_{n_1}$ 是总体 $X\sim N(\mu_1,\sigma_1^2)$ 的一个样本, $Y_1,Y_2,\cdots,Y_{n_2}$ 是总体 $Y\sim N(\mu_2,\sigma_2^2)$ 的一个样本,两个样本相互独立. 令

$$\hat{\sigma}_1^2=\frac{1}{n_1}\sum_{i=1}^{n_1}(X_i-\mu_1)^2,\quad \hat{\sigma}_2^2=\frac{1}{n_2}\sum_{i=1}^{n_2}(Y_i-\mu_2)^2,$$

求 $F=\frac{\hat{\sigma}_1^2}{\hat{\sigma}_2^2}$ 的抽样分布(μ_1,μ_2 已知).

解 因为 $$\chi_1^2=\frac{1}{\sigma^2}\sum_{i=1}^{n_1}(X_i-\mu_1)^2\sim\chi^2(n_1),$$

$$\chi_2^2=\frac{1}{\sigma^2}\sum_{i=1}^{n_2}(Y_i-\mu_2)^2\sim\chi^2(n_2),$$

χ_1^2 与 χ_2^2 相互独立，由 F 分布定义知

$$F=\frac{\chi_1^2}{n_1}\Big/\frac{\chi_2^2}{n_2}=\frac{\hat{\sigma}_1^2}{\hat{\sigma}_2^2}\sim F(n_1,n_2).$$

我们一般只习惯于总体 $X\sim N(0,1)$ 下的 χ^2 分布，而不习惯总体 $X\sim N(\mu,\sigma^2)$ 下的 χ^2 分布. 必须适应这种转换.

例 4 设 $X_1,X_2,\cdots,X_n$ 是总体 $X\sim N(\mu,\sigma^2)$ 的一个样本，μ,σ^2 均未知，则下列结论（ ）正确.

(A) $S^2=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2\sim\chi^2(n-1)$；

(B) $\hat{\sigma}^2=\frac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})^2\sim\chi^2(n-1)$；

(C) $\frac{1}{\sigma^2}\sum_{i=1}^{n}(X_i-\overline{X})^2\sim\chi^2(n-1)$；

(D) $\frac{1}{\sigma^2}\sum_{i=1}^{n}(X_i-\overline{X})^2\sim\chi^2(n)$.

解 选(C). 同上例，当 $X\sim N(\mu,\sigma^2)$ 时，有 $\frac{1}{\sigma^2}\sum_{i=1}^{n}(X_i-\mu)^2\sim\chi^2(n)$，但 μ,σ^2 未知. 以 $\overline{X}$ 代替 μ，则 $\overline{X}$ 为一个约束条件，所以 $\frac{1}{\sigma^2}\sum_{i=1}^{n}(X_i-\overline{X})^2\sim\chi^2(n-1)$.

例 5 设 $X_1,X_2,\cdots,X_n$ 为 $X\sim N(0,1)$ 的一个样本，$\overline{X}$ 与 S^2 为样本方差与样本均值，则（ ）成立.

(A) $\overline{X}\sim N(0,1)$； (B) $n\overline{X}\sim N(0,1)$；

(C) $\sum_{i=1}^{n}X_i^2\sim\chi^2(n)$； (D) $\overline{X}/S\sim t(n-1)$.

解 选(B)、(C). 因为 $\overline{X}\sim N(0,1/n)$，所以(A)不正确，(B)正

确.

因为 $X_1,X_2,\cdots,X_n$ 相互独立，$X_i\sim N(0,1)$，所以 $\sum_{i=1}^{n}X_i^2\sim\chi^2(n)$，(C)正确.

当 $\mu=0$ 时，知 $\frac{\overline{X}-\mu}{S/\sqrt{n}}=\frac{\overline{X}}{S/\sqrt{n}}\sim t(n-1)$，但 $n\neq1$ 时，$\frac{\overline{X}}{S/\sqrt{n}}\neq\frac{\overline{X}}{S}$，所以(D)不正确.

例6 设 $X_1,X_2,\cdots,X_5$ 是总体 $X\sim N(0,1)$ 的一个样本，若统计量 $U=c(X_1+X_2)\Big/\sqrt{X_3^2+X_4^2+X_5^2}\sim t(n)$，试确定 c 与 n.

解 因为 X_i $(i=1,2,\cdots,5)$ 相互独立且同分布，所以

$$(X_1+X_2)/\sqrt{2}\sim N(0,1),\quad X_3^2+X_4^2+X_5^2\sim\chi^2(3),$$

且两者相互独立. 由 t 分布定义知

$$U=\frac{X_1+X_2}{\sqrt{2}}\Big/\sqrt{(X_3^2+X_4^2+X_5^2)/3}\sim t(3),$$

故可确定 $c=\sqrt{3/2}$，$n=3$.

例7 设 $X_1,X_2,\cdots,X_n$ 为总体 $X\sim N(\mu,\sigma^2)$ 的一个样本，$\overline{X}$ 和 S^2 为样本均值和样本方差. 又设新增加一个试验量 X_{n+1}，X_{n+1} 与 $X_1,X_2,\cdots,X_n$ 也相互独立，求统计量 $U=\frac{X_{n+1}-\overline{X}}{S}\sqrt{\frac{n}{n+1}}$ 的分布.

解 因为 $\quad\overline{X}\sim N(\mu,\sigma^2/n),\quad X_{n+1}\sim N(\mu,\sigma^2)$，

所以 $\quad(X_{n+1}-\overline{X})\sim N(0,(n+1)\sigma^2/n)$，

于是 $\quad(X_{n+1}-\overline{X})\Big/\left(\sigma\sqrt{\frac{n+1}{n}}\right)\sim N(0,1).$

又 $\frac{(n-1)S^2}{\sigma^2}\sim\chi^2(n-1)$，且 S^2 与 $(X_{n+1}-\overline{X})$ 相互独立. 由 t 分布定义，有

$$U=\frac{X_{n+1}-\overline{X}}{\sigma\sqrt{(n+1)/n}}\Big/\sqrt{\frac{(n-1)S^2}{\sigma^2(n-1)}}$$

$$=\frac{X_{n+1}-\overline{X}}{S}\sqrt{\frac{n}{n+1}}\sim t(n-1).$$

例8 设$X\sim t(k)$,问:$Y=X^2$服从什么分布?并确定其参数.

解 因为$X\sim t(k)$,依t分布定义,$X=\dfrac{U}{\sqrt{V/k}}$,其中$U\sim N(0,1)$,$V\sim\chi^2(k)$,且U,V相互独立.

又由$U\sim N(0,1)$知,$U^2\sim\chi^2(1)$,且U^2与V也相互独立,于是

$$X^2=(U^2/1)/(V/k)\sim F(1,k),$$

即$Y=X^2$服从F分布,参数为$(1,k)$.

例9 设X_1,X_2为总体$X\sim N(0,\sigma^2)$的一个样本,问:$Y=(X_1+X_2)^2/(X_1-X_2)^2$服从什么分布?并确定其参数.

解 因为$X\sim N(0,\sigma^2)$,X_1,X_2相互独立,所以(X_1+X_2)和(X_1-X_2)都服从$N(0,2\sigma^2)$,且

$$[(X_1+X_2)/(\sqrt{2}\sigma)]^2\sim\chi^2(1),$$

$$[(X_1-X_2)/(\sqrt{2}\sigma)]^2\sim\chi^2(1),$$

从而,由F分布定义知

$$Y=\frac{[(X_1+X_2)/(\sqrt{2}\sigma)]^2}{[(X_1-X_2)/(\sqrt{2}\sigma)]^2}=\frac{(X_1+X_2)^2}{(X_1-X_2)^2}\sim F(1,1),$$

即Y服从F分布,参数为$(1,1)$.

例10 从总体$X\sim N(\mu,\sigma^2)$中抽取容量为16的样本.在下列情形下分别求$\overline{x}$与μ之差的绝对值小于2的概率:

(1) 已知$\sigma^2=25$; (2) σ^2未知,但$S^2=20.8$.

解 (1) 由$\sigma=5$,统计量$U=(\overline{X}-\mu)\Big/\dfrac{\sigma}{\sqrt{n}}\sim N(0,1)$,有

$$\begin{aligned}P\{|\overline{x}-\mu|<2\}&=P\left\{|\overline{x}-\mu|\Big/\frac{\sigma}{\sqrt{n}}<2\Big/\frac{5}{\sqrt{16}}\right\}\\&=P\{|u|<1.6\}=2\Phi(1.6)-1\\&=0.8904.\end{aligned}$$

(2) 由统计量 $T=|\bar{x}-\mu|\Big/\frac{S}{\sqrt{n}}\sim t(n-1)$，有

$$P\{|\bar{x}-\mu|<2\}=P\left\{|\bar{x}-\mu|\Big/\frac{S}{\sqrt{n}}<2\Big/\frac{\sqrt{20.8}}{\sqrt{16}}\right\}$$

$$=P\{|t|<1.76\}=1-2\times 0.05=0.90.$$

例 11 设 $X_1,X_2,\cdots,X_{10}$ 是总体 $X\sim N(\mu,0.5^2)$ 的一个样本，

(1) 已知 $\mu=0$，求 $P\left\{\sum\limits_{i=1}^{n}X_i^2\geqslant 4\right\}$；

(2) μ 未知，求 $P\left\{\sum\limits_{i=1}^{n}(X_i-\overline{X})^2\geqslant 2.85\right\}$.

解 (1) $\mu=0$ 时，$\chi_1^2=\frac{1}{\sigma^2}\sum\limits_{i=1}^{10}X_i^2\sim\chi^2(10)$，于是

$$P\left\{\sum_{i=1}^{10}X_i^2\geqslant 4\right\}=P\left\{\frac{1}{\sigma^2}\sum_{i=1}^{10}X_i\geqslant\frac{4}{0.5^2}\right\}$$

$$=P\{\chi_1^2\geqslant 16\}\xlongequal{\text{查表}}0.10.$$

(2) μ 未知时，$\chi_2^2=\frac{1}{\sigma^2}\sum\limits_{i=1}^{10}(X_i-\overline{X})^2\sim\chi^2(9)$，于是

$$P\left\{\sum_{i=1}^{10}(X_i-\overline{X})^2\geqslant 2.85\right\}=P\left\{\frac{1}{\sigma^2}\sum_{i=1}^{10}(X_i-\overline{X})^2\geqslant\frac{2.85}{0.5^2}\right\}$$

$$=P\{\chi_2^2\geqslant 11.4\}\xlongequal{\text{查表}}0.25.$$

例 12 设 $X_1,X_2,\cdots,X_n$ 是总体 $X\sim N(\mu,\sigma^2)$ 的一个样本，S^2 是样本方差. 试确定 n 多大时，有 $P\{S^2/\sigma^2\leqslant 1.5\}\geqslant 0.95$.

解 因为 $$\frac{(n-1)S^2}{\sigma^2}\sim\chi^2(n-1),$$

要 $$P\left\{\frac{S^2}{\sigma^2}\leqslant 1.5\right\}=P\left\{\frac{(n-1)S^2}{\sigma^2}\leqslant 1.5(n-1)\right\}\geqslant 0.95,$$

即要 $P\left\{\frac{(n-1)S^2}{\sigma^2}>1.5(n-1)\right\}\leqslant 0.05$. 查表知

$$1.5\times(27-1)=39>\chi_{0.05}^2(26)=38.885,$$

$$1.5\times(26-1)=37.5<\chi_{0.05}^2(25)=37.652,$$

于是，取 $n=27$.

例13 设 $X_1,X_2,\cdots,X_{20}$ 为总体 $X\sim N(\mu,\sigma^2)$ 的一个样本，求：

(1) $P\left\{10.9\leqslant\frac{1}{\sigma^2}\sum_{i=1}^{20}(X_i-\mu)^2\leqslant 37.6\right\}$；

(2) $P\left\{12.4\leqslant\frac{1}{\sigma^2}\sum_{i=1}^{20}(X_i-\mu)^2\leqslant 40\right\}$.

解 因为 $X\sim N(\mu,\sigma^2)$，所以 $\chi^2=\frac{1}{\sigma^2}\sum_{i=1}^{20}(X_i-\mu)^2\sim\chi^2(20)$.

$$
\begin{aligned}
(1)\ &P\left\{10.9\leqslant\frac{1}{\sigma^2}\sum_{i=1}^{20}(X_i-\mu)^2\leqslant 37.6\right\}\\
&=P\{10.9\leqslant\chi^2\leqslant 37.6\}\\
&=P\{\chi^2\leqslant 37.6\}-P\{\chi^2\leqslant 10.9\}\\
&=1-P\{\chi^2>37.6\}-[1-P\{\chi^2>10.9\}]\\
&=P\{\chi^2>10.9\}-P\{\chi^2>37.6\}\\
&\overset{\text{查表}}{=\!=}0.95-0.01=0.94.
\end{aligned}
$$

查表方法是：在 χ^2 分布表中先查到 $n=20$ 的一行，再横向查得与10.9接近的10.851，该列对应的 $\alpha=0.95$ 即为所求概率.

同理查得 $P\{\chi^2>37.6\}=0.01$.

(2) 类似地，有

$$
\begin{aligned}
&P\left\{12.4\leqslant\frac{1}{\sigma^2}\sum_{i=1}^{20}(X_i-\mu)^2\leqslant 40\right\}\\
&=P\{\chi^2>12.4\}-P\{\chi^2>40\}\\
&=0.90-0.005=0.895.
\end{aligned}
$$

例14 设 $X_1,X_2,\cdots,X_{16}$ 是总体 $X\sim N(\mu,\sigma^2)$ 的一个样本，μ,σ^2 为未知，而 $\bar{x}=12.5$，$S^2=5.333$，求 $P\{|\bar{x}-\mu|<0.4\}$.

解 因为 σ 未知，所以有

$$t=\frac{\bar{x}-\mu}{S/\sqrt{n}}\sim t(n-1).$$

将 $n=16$，$S=\sqrt{5.333}=2.309$ 代入，得 $t=\frac{\bar{x}-\mu}{0.5773}\sim t(15)$.

$$P\{|\bar{x}-\mu|<0.4\}=P\left\{\frac{|\bar{x}-\mu|}{0.5773}<\frac{0.4}{0.5773}\right\}=P\{|t|<0.692\}$$
$$=1-P\{t\geqslant 0.692\}-P\{t<0.692\}$$
$$=1-2P\{t\geqslant 0.692\}=1-2\times 0.25=0.5.$$

查 t 分布表，方法是：先查到 $n=16-1=15$ 的一行，横向查到 0.692，对应的 α 即为概率 $P\{t\geqslant 0.692\}$.

例 15 设 $X_1, X_2, \cdots, X_9$ 为总体 $X \sim N(\mu, 2^2)$ 的一个样本，若记 $Y=\sum_{i=1}^{9}(X_i-\overline{X})^2$，求满足 $P\{Y\geqslant \alpha_2\}=P\{Y\leqslant \alpha_1\}=0.05$ 的 α_1 和 α_2.

解 因为 $D(X)=4$，所以

$$U=\frac{Y}{4}=\frac{1}{4}\sum_{i=1}^{9}(X_i-\overline{X})^2 \sim \chi^2(8).$$

故
$$P\{Y\geqslant \alpha_2\}=P\{U\geqslant \alpha_2/4\}=0.05.$$

查表知，若 $P\{U\geqslant \chi^2_{0.05}(8)\}=0.05$，则 $\alpha_2/4=\chi^2_{0.05}(8)=15.507$，于是 $\alpha_2=62.028$.

类似地，由 $P\{Y\leqslant \alpha_1\}=P\{U\leqslant \alpha_1/4\}=0.05$，查表知 $\alpha_1/4=\chi^2_{0.95}(8)=2.733$，于是 $\alpha_1=10.932$.

例 16 设总体 X 服从指数分布，概率密度

$$f(x)=\begin{cases}\frac{1}{\theta}e^{-x/\theta}, & x>0, \theta>0, \\ 0, & \text{其它},\end{cases}$$

$X_1, X_2, \cdots, X_n$ 为 X 的一个样本，证明：$\frac{2n\overline{X}}{\theta}\sim \chi^2(2n)$.

证一 这是一种简捷证法. 因为 $\chi^2(2)$ 的概率密度为

$$f(y)=\begin{cases}\frac{1}{2}e^{-y/2}, & y>0, \\ 0, & \text{其它},\end{cases}$$

所以，$\chi^2(2)$ 分布也可以看作 $\theta=2$ 的指数分布. 令 $Y=2X/\theta$，则由 χ^2 分布的可加性，有

$$\frac{2n\overline{X}}{\theta}=\frac{2}{\theta}\sum_{i=1}^{n}X_i=\sum_{i=1}^{n}\frac{2X_i}{\theta}=\sum_{i=1}^{n}Y_i\sim\chi^2(2n).$$

证二 因为 $X\sim e(\theta)$，即 $X\sim\Gamma(1,1/\theta)$，由于参数为 α,β 的 Γ 分布的密度函数为

$$f(x)=\begin{cases}\frac{\beta}{\Gamma(\alpha)}(\beta x)^{\alpha-1}e^{-\beta x}, & x>0,\\ 0, & x\leqslant 0.\end{cases}$$

由 Γ 分布的参数可加性知，$Y=\sum_{i=1}^{n}X_i\sim\Gamma\left(n,\frac{1}{\theta}\right)$. 而

$$Z=\frac{2n\overline{X}}{\theta}=\frac{2}{\theta}\sum_{i=1}^{n}X_i=\frac{2}{\theta}Y,$$

由随机变量函数分布的定理(公式法)知

$$f(Z)=\begin{cases}\frac{1/\theta}{\Gamma(n)}\left(\beta\frac{Z\theta}{2}\right)^{n-1}e^{-\frac{1}{\theta}\frac{Z\theta}{2}}\cdot\frac{\theta}{2}, & Z>0,\\ 0, & Z\leqslant 0,\end{cases}$$

即 $Z=\frac{2n\overline{X}}{\theta}$ 的密度函数为

$$f(Z)=\begin{cases}\frac{1}{2^n\Gamma(n)}Z^{n-1}\cdot e^{-Z/2}, & Z>0,\\ 0, & Z\leqslant 0.\end{cases}$$

此式正好是 $\chi^2(2n)$ 分布的密度函数，于是证得

$$\frac{2n\overline{X}}{\theta}\sim\chi^2(2n).$$

这两种证法实质上是一样的.

例17 设 $X_1,X_2,\cdots,X_{10}$ 是总体 $X\sim N(0,1)$ 的一个样本，$\overline{X}$ 和 S^2 分别是样本均值和样本方差. 令 $Y=10\overline{X}^2/S^2$，若有 $P\{Y>\lambda\}=0.01$，则 λ 应为多少？

解 由 t 分布定义知 $\overline{X}\sim N\left(0,\frac{1}{10}\right)$，而 $\overline{X}\Big/\frac{S}{\sqrt{10}}\sim t(9)$，于是

$$Y=T^2=\frac{10\overline{X}^2}{S^2}\sim F(1,9).$$

查 F 分布表知 $\lambda=F_{0.01}(1,9)=10.56$.

例 18 设 $X_1,X_2,\cdots,X_8$ 是总体 $X\sim N(\mu,20)$ 的一个样本，Y_1，$Y_2,\cdots,Y_{10}$ 是 $Y\sim N(\mu,35)$ 的一个样本，X 与 Y 相互独立，S_1^2 和 S_2^2 是各自的样本方差，求 $P\{S_1^2\geqslant 2S_2^2\}$.

解 因为

$$F=\frac{S_1^2/20}{S_2^2/35}=1.75\frac{S_1^2}{S_2^2}\sim F(8-1,10-1)=F(7,9),$$

所以

$$P\{S_1^2\geqslant 2S_2^2\}=P\left\{\frac{S_1^2}{S_2^2}\geqslant 2\right\}=P\left\{\frac{S_1^2/20}{S_2^2/35}\geqslant 2\times\frac{35}{20}\right\}=P\{F\geqslant 3.5\}.$$

查 F 分布表知 $F_{0.05}(7,9)=3.29$，$F_{0.025}(7,9)=4.20$，而 $3.29<3.5<4.20$，于是

$$0.025\leqslant P\{S_1^2\geqslant 2S_2^2\}\leqslant 0.05.$$

例 19 设 $X_1,X_2,\cdots,X_{10}$ 是 $X\sim N(10,2^2)$ 的一个样本，Y_1,Y_2，$\cdots,Y_5$ 是 $Y\sim N(20,2^2)$ 的一个样本，两者相互独立. 令

$$F_1=\frac{\sum_{i=1}^{10}(X_i-10)^2}{\sum_{i=1}^{5}(Y_i-20)^2},\quad F_2=\frac{\sum_{i=1}^{5}(Y_i-\overline{Y})^2}{\sum_{i=1}^{10}(X_i-\overline{X})^2},$$

(1) 已知 $P\{F_1\leqslant\alpha_1\}=0.05$，求 α_1；

(2) 已知 $P\{F_2\leqslant\alpha_2\}=0.01$，求 α_2.

解 (1) 此时，$F_1\sim F(5,5)$，$P\{F_1\leqslant\alpha_1\}=P\left\{\frac{1}{F_1}\geqslant\frac{1}{\alpha_1}\right\}=0.05$. 查 F 分布表知

$$\frac{1}{\alpha_1}=F_{0.05}(5,5)=5.05\Longrightarrow\alpha_1=0.198.$$

(2) 因为

$$F=\frac{1}{9}\sum_{i=1}^{10}(X_i-\overline{X})^2\Big/\left[\frac{1}{4}\sum_{i=1}^{5}(Y_i-\overline{Y})^2\right]\sim F(9,4),$$

所以

$$P\{F_2\leqslant\alpha_2\}=P\left\{\sum_{i=1}^{5}(Y_i-\overline{Y})^2\Big/\sum_{i=1}^{10}(X_i-\overline{X})^2\leqslant\alpha_2\right\}$$
$$=P\left\{\sum_{i=1}^{10}(X_i-\overline{X})^2\Big/\sum_{i=1}^{5}(Y_i-\overline{Y})^2\geqslant\frac{1}{\alpha_2}\right\}$$
$$=P\left\{\frac{1}{9}\sum_{i=1}^{10}(X_i-\overline{X})^2\Big/\left[\frac{1}{4}\sum_{i=1}^{5}(Y_i-\overline{Y})^2\right]\geqslant\frac{4}{9\alpha_2}\right\}$$
$$=P\left\{F\geqslant\frac{4}{9\alpha_2}\right\}.$$

查 F 分布表知

$$\frac{4}{9\alpha_2}=F_{0.01}(9,4)=14.66\Longrightarrow\alpha_2=0.03.$$

例 20 设 $X_1,X_2,\cdots,X_n$ 是总体 $X\sim N(\mu,\sigma^2)$ 的一个样本，试证：$E\left[\sum_{i=1}^{n}(X_i-\overline{X})^2\right]^2=(n^2-1)\sigma^4$.

证 因为 $\chi^2=\sum_{i=1}^{n}(X_i-\overline{X})^2/\sigma^2\sim\chi^2(n-1)$，$E(\chi^2)=n-1$，$D(\chi^2)=2(n-1)$，所以

$$E\left[\sum_{i=1}^{n}(X_i-\overline{X})^2\right]^2=\sigma^4E\left[\sum_{i=1}^{n}(X_i-\overline{X})^2/\sigma^2\right]^2$$
$$=\sigma^4E(\chi^2)^2=\sigma^4[D(\chi^2)+[E(\chi^2)]^2]$$
$$=\sigma^4[2(n-1)+(n-1)^2]=(n^2-1)\sigma^4.$$

例 21 设 X_1,X_2 是总体 $X\sim N(\mu,\sigma^2)$ 的一个样本，证明：X_1+X_2 与 X_1-X_2 相互独立.

证 因为

$$\mathrm{cov}(X_1+X_2,X_1-X_2)$$
$$=E(X_1+X_2)(X_1-X_2)-E(X_1+X_2)E(X_1-X_2)$$
$$=E(X_1^2-X_2^2)-[E(X_1)+E(X_2)][E(X_1)-E(X_2)],$$

而 $$E(X_1^2)=E(X_2^2),\quad E(X_1)=E(X_2),$$

故 $$\mathrm{cov}(X_1+X_2,X_1-X_2)=0.$$

又 $X_1+X_2\sim N(2\mu,2\sigma^2)$，$X_1-X_2\sim N(0,2\sigma^2)$，两个正态总体

不相关则一定相互独立，所以，由 $\mathrm{cov}(X_1+X_2,X_1-X_2)=0$ 知，X_1+X_2 与 X_1-X_2 不相关，必相互独立.

例 22 设随机变量 $X\sim F(m,m)$，证明：

$$P\{X\leqslant 1\}=P\{X\geqslant 1\}=0.5.$$

证 若 $X\sim F(m,n)$，则 $1/X\sim F(n,m)$. 由于 $m=n$，故 X 与 $1/X$ 服从同一分布，于是

$$P\{X\leqslant 1\}=P\{1/X\leqslant 1\}=P\{X\geqslant 1\}.$$

而
$$P\{X\leqslant 1\}+P\{X\geqslant 1\}=1,$$

因此
$$P\{X\leqslant 1\}=P\{X\geqslant 1\}=0.5.$$

例 23 设 $X_1,X_2,\cdots,X_{n_1}$ 是 $X\sim N(\mu_1,\sigma^2)$ 的一个样本，$Y_1,Y_2,\cdots,Y_{n_2}$ 是 $Y\sim N(\mu_2,\sigma^2)$ 的一个样本，两者相互独立. $\overline{X},\overline{Y}$ 是它们的样本均值，S_1^2,S_2^2 是它们的样本方差，c,d 是常数. 证明：

$$t=\frac{c(\overline{X}-\mu_1)+d(\overline{Y}-\mu_2)}{S_W\sqrt{c^2/n_1+d^2/n_2}}\sim t(n_1+n_2-2),$$

其中
$$S_W^2=[(n_1-1)S_1^2+(n_2-1)S_2^2]/(n_1+n_2-2).$$

证 因为 $E(c\overline{X}+d\overline{Y})=cE(\overline{X})+dE(\overline{Y})=c\mu_1+d\mu_2,$

$$D(c\overline{X}+d\overline{Y})=c^2D(\overline{X})+d^2D(\overline{Y})=\sigma^2(c^2/n_1+d^2/n_2),$$

所以
$$U=\frac{c(\overline{X}-\mu_1)+d(\overline{Y}-\mu_2)}{\sigma\sqrt{c^2/n_1+d^2/n_2}}$$
$$=\frac{(c\,\overline{X}+d\,\overline{Y})+(c\mu_1+d\mu_2)}{\sigma\sqrt{c^2/n_1+d^2/n_2}}\sim N(0,1).$$

又知 $(n_1-1)S_1^2/\sigma^2\sim\chi^2(n_1-1)$， $(n_2-1)S_2^2/\sigma^2\sim\chi^2(n_2-1)$.

它们相互独立，由 χ^2 分布的可加性知

$$\chi^2=\frac{1}{\sigma^2}[(n_1-1)S_1^2+(n_2-1)S_2^2]\sim\chi^2(n_1+n_2-2).$$

而 U 与 χ^2 相互独立，依 t 分布定义，有

$$t=\frac{c(\overline{X}-\mu_1)+d(\overline{Y}-\mu_2)}{S_W\sqrt{c^2/n_1+d^2/n_2}}\sim t(n_1+n_2-2).$$

例 24 设$X\sim F(k_1,k_2)$，证明：$\frac{1}{X}\sim F(k_2,k_1)$，从而

$$F_{1-\alpha}(k_1,k_2)=1/F_\alpha(k_2,k_1).$$

证 令$X=\frac{U}{k_1}\Big/\frac{V}{k_2}$，设$U\sim\chi^2(k_1)$，$V\sim\chi^2(k_2)$，$U,V$相互独立，则

$$\frac{1}{X}=\frac{V}{k_2}\Big/\frac{U}{k_1}\sim F(k_2,k_1).$$

由 $P\left\{\frac{1}{X}\geqslant F_\alpha(k_2,k_1)\right\}=\alpha\Longrightarrow P\left\{\frac{1}{X}\leqslant F_\alpha(k_2,k_1)\right\}=1-\alpha,$

即 $$P\{X\geqslant 1/F_\alpha(k_2,k_1)\}=1-\alpha.$$

因为 $$P\{X\geqslant F_{1-\alpha}(k_1,k_2)\}=1-\alpha,$$

故 $$F_{1-\alpha}(k_1,k_2)=1/F_\alpha(k_2,k_1).$$

硕士研究生入学试题分析

一、本章考试要求

1. 理解总体、简单随机样本、统计量、样本均值、样本方差及样本矩的概念，了解经验分布函数.

2. 了解χ^2分布、t分布和F分布的定义及性质，了解分位数的概念并会查表计算.

3. 了解正态总体的某些常用抽样分布.

二、本章的重点内容

从已知总体中抽取一个随机样本，构造一个统计量，然后确定这个统计量的分布，并确定其参数或者自由度.

数理统计的基本概念.

1. 设$X_1,X_2,\cdots,X_n$ $(n\geqslant 2)$为来自总体$N(0,1)$的简单随机样本，$\overline{X}$为样本均值，S^2为样本方差，则（ ）.

(A) $n\overline{X}\sim N(0,1)$； (B) $nS^2\sim\chi^2(n)$；

(C) $\dfrac{(n-1)\overline{X}}{S}\sim t(n-1)$； (D) $\dfrac{n-1}{\sum\limits_{i=2}^{n}X_i^2}\sim F(1,n-1)$.

（2005 年一）

解 选(D). 因为$n\overline{X}=X_1+X_2+\cdots+X_n$,所以(A)不成立.
因为$(n-1)S^2/\sigma^2\sim\chi^2(n-1)$,所以(B)不成立.
因为$(\overline{X}-\mu)/(S/\sqrt{n})\sim t(n-1)$,所以(C)不成立.
而$X_1^2\sim\chi^2(1)$, $\sum\limits_{i=2}^{n}X_i^2\sim\chi^2(n-1)$,故(D)成立.

2. 设随机变量$X\sim t(n)$ $(n>1)$, $Y=\dfrac{1}{X^2}$,则(　　).

(A) $Y\sim\chi^2(n)$； (B) $Y\sim\chi^2(n-1)$；

(C) $Y\sim F(n,1)$； (D) $Y\sim F(1,n)$.

（2003 年一）

解 选(C). 因为$X\sim t(n)$,所以$X=U/\sqrt{V/n}$,其中$U\sim N(0,1)$, $V\sim\chi^2(n)$. 由$Y=1/X^2$知,$Y=(V/n)/U^2$,则由$V\sim\chi^2(n)$, $U^2\sim\chi^2(1)$得$Y\sim F(n,1)$.

3. 设总体X服从参数为2的指数分布,$X_1,X_2,\cdots,X_n$为来自总体X的简单随机样本,则当$n\to\infty$时,$Y_n=\dfrac{1}{n}\sum\limits_{i=1}^{n}X_i^2$依概率收敛于________.

（2003 年三）

解 因为$E(X_i)=1/2$, $D(X_i)=1/4$, $i=1,2,\cdots,n$,所以样本二阶矩依概率收敛于总体矩,即$Y_n=\dfrac{1}{n}\sum\limits_{i=1}^{n}X_i^2$收敛于1/2.

4. 设$X_1,X_2,\cdots,X_n$是来自正态总体$N(\mu,\sigma^2)$的简单随机样本,$\overline{X}$是样本均值,记

$$S_1^2=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2,\quad S_2^2=\frac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})^2,$$

$$S_3^2=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\mu)^2,\quad S_4^2=\frac{1}{n}\sum_{i=1}^{n}(X_i-\mu)^2,$$

则服从自由度为 $n-1$ 的 t 分布的随机变量是(　　).

(A) $t=\dfrac{\overline{X}-\mu}{S_1/\sqrt{n-1}}$；　　(B) $t=\dfrac{\overline{X}-\mu}{S_2/\sqrt{n-1}}$；

(C) $t=\dfrac{\overline{X}-\mu}{S_3/\sqrt{n-1}}$；　　(D) $t=\dfrac{\overline{X}-\mu}{S_4/\sqrt{n-1}}$.　　(1994 年四)

解　选(B). 由定理

$$(\overline{X}-\mu)/(S_1/\sqrt{n})\sim t(n-1),$$

而
$$S_1/\sqrt{n}=\frac{1}{\sqrt{n-1}}\sum_{i=1}^{n}(X_i-\overline{X})^2/\sqrt{n}$$
$$=\frac{1}{\sqrt{n}}\sum_{i=1}^{n}(X_i-\overline{X})^2/\sqrt{n-1}=S_2/\sqrt{n-1},$$

故
$$t=(\overline{X}-\mu)/(S_2/\sqrt{n-1}).$$

5. 设随机变量 X 和 Y 相互独立，且都服从正态分布 $N(0,3^2)$，而 $X_1,X_2,\cdots,X_9$ 和 $Y_1,Y_2,\cdots,Y_9$ 分别是来自总体 X 和 Y 的简单随机样本，则统计量

$$U=(X_1+X_2+\cdots+X_9)/\sqrt{Y_1^2+Y_2^2+\cdots+Y_9^2}$$

服从________分布，参数为________.　　(1997 年三)

解　因为 $X_1+X_2+\cdots+X_9\sim N(0,1)$，$Y_1^2+Y_2^2+\cdots+Y_9^2\sim\chi^2(9)$，而 $t=X/\sqrt{Y/n}\sim t(n)$，故 $U\sim t(9)$.

6. 设 X_1,X_2,X_3,X_4 是来自正态总体 $N(0,2^2)$ 的简单随机样本，$X=a(X_1-2X_2)^2+b(3X_3-4X_4)^2$，则当 $a=1/20$，$b=1/100$ 时，统计量 X 服从________分布，且自由度为________.

(1998 年三)

解　令 $X_1'=X_1-2X_2$，$X_2'=3X_3-4X_4$，则 $X_1'\sim N(0,20)$，$X_2'\sim N(0,100)$. 当 $a=1/20$ 时，$\sqrt{a}\,X_1'\sim N(0,1)$；当 $b=1/100$ 时，$\sqrt{b}\,X_2'\sim N(0,1)$. 所以 $X\sim\chi^2(2)$.

7. 设 $X_1,X_2,\cdots,X_9$ 是来自正态总体的简单随机样本，

$$Y_1=(X_1+X_2+\cdots+X_6)/6,\quad Y_2=(X_7+X_8+X_9)/3,$$

$$S^2=\frac{1}{2}\sum_{i=7}^{9}(X_i-Y_2)^2,\quad Z^2=\sqrt{2}(Y_1-Y_2)/S,$$

证明：统计量 Z 服从自由度为 2 的 t 分布. (1999 年三)

证 $D(X)=\sigma^2$ 为未知，而

$$E(Y_1)=E(Y_2),\quad D(Y_1)=\sigma^2/6,\quad D(Y_2)=\sigma^2/3.$$

由 Y_1 与 Y_2 的独立性知

$$E(Y_1-Y_2)=0,\quad D(Y_1-Y_2)=\sigma^2/6+\sigma^2/3=\sigma^2/2,$$

故 $$U=(Y_1-Y_2)/(\sigma/\sqrt{2})\sim N(0,1).$$

由正态总体样本方差的性质知，$\chi^2=2S^2/\sigma^2\sim\chi^2(n)$.

又由 Y_1 与 Y_2 相互独立知，Y_1 与 S^2 相互独立，Y_2 与 S^2 相互独立，于是 Y_1-Y_2 也与 S^2 相互独立. 从而，由 t 分布随机变量的构造知

$$Z=\sqrt{2}(Y_1-Y_2)/S=U/\sqrt{\chi^2/2}\sim t(2).$$

8. 设总体 X 服从正态分布 $N(0,2^2)$，而 $X_1,X_2,\cdots,X_{15}$ 是来自总体 X 的简单随机样本，则随机变量

$$Y=\frac{X_1^2+X_2^2+\cdots+X_{10}^2}{2(X_{11}^2+X_{12}^2+\cdots+X_{15}^2)}$$

服从________分布，参数为________. (2001 年三)

解 随机变量 $X_1^2+X_2^2+\cdots+X_{10}^2\sim\chi^2(10)$，随机变量 $X_{11}^2+X_{12}^2+\cdots+X_{15}^2\sim\chi^2(5)$，所以

$$\frac{X/n_1}{Z/n_2}=\frac{(X_1^2+X_2^2+\cdots+X_{10}^2)/10}{(X_{11}^2+X_{12}^2+\cdots+X_{15}^2)/5}=\frac{X_1^2+X_2^2+\cdots+X_{10}^2}{2(X_{11}^2+X_{12}^2+\cdots+X_{15}^2)}\sim F(10,5).$$

9. 设随机变量 X 和 Y 都服从标准正态分布，则(　　).

(A) $X+Y$ 服从正态分布；　(B) X^2+Y^2 服从 χ^2 分布；

(C) X^2 和 Y^2 都服从 χ^2 分布；　(D) X^2/Y^2 服从 F 分布.

(2002 年三)

解 选(C). 因为 X 和 Y 不一定独立，X^2+Y^2 也不一定独立，所以(A)、(B)、(D)不一定能成立.

第七章　参数估计

第一节　点　估　计

主要内容

当总体 X 的分布形式已知，但它的一个或多个参数未知时，用总体 X 的一个样本来估计参数的真值，称为参数的点估计.

设总体 X 的分布函数 $F(x,\theta)$ 已知，θ 未知，$X_1,X_2,\cdots,X_n$ 为来自总体 X 的一个样本，$x_1,x_2,\cdots,x_n$ 为一组样本观察值. 点估计就是要构造一个适当的统计量 $\hat{\theta}(X_1,X_2,\cdots,X_n)$，用它的一个观察值 $\hat{\theta}(x_1,x_2,\cdots,x_n)$ 来估计未知参数 θ 的真值. $\hat{\theta}(X_1,X_2,\cdots,X_n)$ 称为参数 θ 的估计量，是总体 X 的样本 $X_1,X_2,\cdots,X_n$ 的函数；$\hat{\theta}(x_1,x_2,\cdots,x_n)$ 称为 θ 的一个估计值.

1. 矩估计法

用样本矩作为总体矩的估计量，以样本矩的连续函数作为相应总体矩的连续函数的估计量，这种估计方法称为矩估计法.

设总体 X 的分布函数为 $F(x;\theta_1,\theta_2,\cdots,\theta_k)$，其中 $\theta_1,\theta_2,\cdots,\theta_k$ 为未知参数. 若 X 为连续型随机变量，其概率密度为 $f(x;\theta_1,\theta_2,\cdots,\theta_k)$；若 X 为离散型随机变量，其分布律为 $P\{X=x\}=p(x;\theta_1,\theta_2,\cdots,\theta_k)$. 设 $X_1,X_2,\cdots,X_n$ 为 X 的一个样本，且总体 X 的前 k 阶矩

$$\mu_i=E(X^i)=\int_{-\infty}^{+\infty}x^i f(x;\theta_1,\theta_2,\cdots,\theta_k)\mathrm{d}x$$

或 $\mu_i=E(X^i)=\sum_{x\in R_x}x^i p(x;\theta_1,\theta_2,\cdots,\theta_k)\quad(i=1,2,\cdots,k)$

存在(它们也是$\theta_1,\theta_2,\cdots,\theta_k$ 的函数). 又,样本矩

$$A_i=\frac{1}{n}\sum_{j=1}^{n}X_j^i\xrightarrow{p}\mu_i,$$

则令

$$\mu_i=A_i\quad(i=1,2,\cdots,k),$$

得一个含k个未知参数的联立方程组. 方程组的解$\hat{\theta}_1,\hat{\theta}_2,\cdots,\hat{\theta}_k$即为$\theta_1,\theta_2,\cdots,\theta_n$的矩估计量.

2. 极大似然估计法

若总体X为离散型随机变量,分布律$P\{X=k\}=p(x;\theta_1,\theta_2,\cdots,\theta_k)$,$\theta_1,\theta_2,\cdots,\theta_k$为未知参数,则对于来自总体$X$的一个样本$X_1,X_2,\cdots,X_n$,称

$$L_n(\theta_1,\theta_2,\cdots,\theta_k)=\prod_{i=1}^{n}p(x_i;\theta_1,\theta_2,\cdots,\theta_k)$$

为样本的似然函数. 若总体X为连续型随机变量,概率密度为

$$f(x;\theta_1,\theta_2,\cdots,\theta_k),$$

则对于来自总体X的一个样本$X_1,X_2,\cdots,X_n$,称

$$L_n(\theta_1,\theta_2,\cdots,\theta_k)=\prod_{i=1}^{n}f(x_i;\theta_1,\theta_2,\cdots,\theta_k)$$

为样本的似然函数.

固定一组样本观察值$x_1,x_2,\cdots,x_n$,取$\hat{\theta}$,使

$$L(\hat{\theta})=L(\hat{\theta}_1,\hat{\theta}_2,\cdots,\hat{\theta}_k)=\max_{\theta\in\Theta}L(\theta_1,\theta_2,\cdots,\theta_k),$$

则此时的$\hat{\theta}(x_1,x_2,\cdots,x_n)$称为参数$\theta$的极大(最大)似然估计值. $\hat{\theta}(X_1,X_2,\cdots,X_n)$称为参数$\theta$的极大(最大)似然估计量.

3. 估计量的评选标准

(1) 无偏性　设$\hat{\theta}=\hat{\theta}(X_1,X_2,\cdots,X_n)$为未知参数$\theta$的估计量,若有$E(\hat{\theta})=\theta$ $(\theta\in\Theta)$,则称$\hat{\theta}$为θ的无偏估计量.

(2) 有效性　若$\hat{\theta}_1$和$\hat{\theta}_2$都是θ的无偏估计量,若有$D(\hat{\theta}_1)<D(\hat{\theta}_2)$ $(\theta\in\Theta)$,则称$\hat{\theta}_1$比$\hat{\theta}_2$有效.

(3) 一致性　若 $\hat{\theta}_n$ 是 θ 的估计量，如果对任给的 $\varepsilon>0$，有

$$\lim_{n\to\infty}P\{|\hat{\theta}_n(X_1,X_2,\cdots,X_n)-\theta|<\varepsilon\}=1,$$

则称 $\hat{\theta}_n$ 是 θ 的一致（相合）估计量.

疑难解析

1. 矩估计法的基本思想是什么？矩估计量是否是唯一的？

答　格里文科定理指出，当 $n\to\infty$ 时，经验分布函数 $F_n^*(x)$ 关于 x 均匀地依概率收敛于总体分布函数. 这就使得在大样本下可以用 $F_n^*(x)$ 代替 $F(x)$ 研究统计推断问题的理论依据.

当以 $F_n^*(x)$ 代替 $F(x)$ 时，总体的原点矩 $\mu_k=\int_{-\infty}^{+\infty}x^k\mathrm{d}F(x)$ 的估计 $\hat{\mu}_k=\int_{-\infty}^{+\infty}x^k\mathrm{d}F_n^*(x)=\frac{1}{n}\sum_{i=1}^{n}X_i^k=A_k$，恰好是样本的同阶原点矩. 因此，用样本原点矩代替总体矩是可行的，这是矩估计法的基本思想.

在一般情况下，矩估计不是唯一的. 如 X 服从泊松分布 $\pi(\lambda)$，λ 是未知参数时，可以用 $E(X)=\lambda$，即样本一阶原点矩代替总体一阶原点矩，也可以用 $D(X)=\lambda$，即样本二阶中心矩代替总体二阶中心矩.

2. 极大似然估计法的基本思想是什么？它有什么性质？要注意些什么问题？

答　极大似然估计是利用总体 X 的概率分布以及样本提供的信息所建立的求未知参数估计量的一种方法. 它建立在这样一种直观想法的基础上：假定一个随机试验 E 有若干个可能结果 A_1，$A_2,\cdots,A_n$，如果只进行了一次试验，而结果 A_k 出现了，那么有理由认为试验的条件对结果 A_k 的出现有利，即试验 E 出现结果 A_k 的概率最大. 例如，已知一袋中装有黑、白两色球，比例为 9 : 1，但不知哪一种颜色的球多. 现设黑球所占比例为 p，做放回抽样下的两次随机取球试验，每次取一个，结果两次都取得黑球，于是可以认定，

$p=0.9$. 这是因为，当 $p=0.9$ 时，连续取得两个黑球的概率为 $(0.9)^2=0.81$；而当 $p=0.1$ 时，这时概率只有 0.01，显然两次都取得黑色球对 $p=0.9$ 有利.

极大似然法的基本思想是：适当地选取 θ，使样本似然函数 $L(\theta)$ 的值达到最大，也就是使试验得出结果 $X_1=x_1, X_2=x_2, \cdots, X_n=x_n$ 的概率最大.

3. 估计量的三个评选标准各有什么意义？

答 估计量的三个评选标准各有自己的意义，读者要认真理解.

无偏性是指对估计量 $\hat{\theta}$，有 $E(\hat{\theta})=\theta, \theta\in\Theta$. 因为 $E(\hat{\theta})-\theta$ 反映用 $\hat{\theta}$ 作为 θ 的估计时的系统误差，所以要求 $E(\hat{\theta})=\theta$，即要求不存在系统误差. 也就是说，当用 $\hat{\theta}$ 来估计 θ 时，$\hat{\theta}$ 与 θ 真值的偏差不是估计量本身造成的，而是随机误差造成的，所以在多次重复试验下，有 $E(\hat{\theta})=\theta$.

无偏性的优点是易于验证. 但不是每个参数都有无偏估计量，而有时一个参数又可以有多个无偏估计量.

有效性是对同一参数的两个无偏估计量进行比较而产生的一个标准，即若 $\hat{\theta}_1$ 和 $\hat{\theta}_2$ 都是 θ 的无偏估计量，而 $D(\hat{\theta}_1)<D(\hat{\theta}_2)$，则称 $\hat{\theta}_1$ 比 $\hat{\theta}_2$ 更有效，也就是无偏估计应是方差小的为好. 特别是，若有一估计量 $\hat{\theta}_0$，对任一无偏估计量 θ 有 $D(\hat{\theta}_0)\leqslant D(\hat{\theta})$，则称 $\hat{\theta}_0$ 是 θ 的最小方差无偏估计量.

有效性在理论上和直观上都较合理，而且容易验证，所以使用得较多.

一致性是在 $n\to\infty$ 时，有 $\hat{\theta}_n\to\theta$.

一致性有重要的理论意义与实际意义. 当 n 充分大时，$\hat{\theta}$ 十分接近 θ 的真值. 利用大数定理可以证明，常用估计量都满足一致性. 如样本矩 A_k 是 μ_k 的无偏估计，也是 μ_k 的一致估计，只是有时让 $n\to\infty$ 不易做到.

方法、技巧与典型例题分析

矩估计法解题的步骤是：首先根据实际问题确定总体的分布与待估计参数，再计算总体矩与样本矩，令总体矩与同阶样本矩相等，解得矩估计量. 矩估计量操作简便可行，如能注意求总体矩的技巧，将更为简捷.

一、矩估计的求法

例 1 设总体 X 的概率分布为

X	1	2	3
p_k	θ^2	$2\theta(1-\theta)$	$(1-\theta)^2$

其中 θ 为未知参数. 现抽得一个样本 $x_1=1, x_2=2, x_3=1$，求 θ 的矩估计值.

解 先求总体一阶原点矩

$$E(X)=1\times\theta^2+2\times2\theta(1-\theta)+3(1-\theta)^2=3-2\theta,$$

一阶样本矩 $$\bar{x}=\frac{1}{3}(1+2+1)=\frac{4}{3}.$$

由 $$E(X)=\bar{x}, \quad 即 \quad 3-2\theta=4/3$$

得 $$\hat{\theta}=5/6,$$

所以 θ 的矩估计值 $\hat{\theta}=5/6$.

例 2 设总体 X 的概率密度为

$$f(x)=\begin{cases}\theta(\theta+1)x^{\theta-1}(1-x), & 0<x<1, \\ 0, & 其它,\end{cases}$$

求 θ 的矩估计量 $\hat{\theta}$ $(0<\theta)$.

解 设 $X_1, X_2, \cdots, X_n$ 是 X 的一个样本，$\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$，则

$$E(X)=\int_0^1 x\theta(\theta+1)x^{\theta-1}(1-x)\mathrm{d}x$$

$$=\theta(\theta+1)\int_0^1(x^\theta-x^{\theta+1})\mathrm{d}x=\frac{\theta}{\theta+2},$$

所以 $$\theta=2\mu_1/(1-\mu_1).$$

由 $\mu_1=A_1$，得

$$\hat{\theta}=2A_1/(1-A_1)=2\overline{X}/(1-\overline{X}).$$

例3 设总体X在$[a,b]$上服从均匀分布，a,b均未知，求a和b的矩估计量.

解 设$X_1,X_2,\cdots,X_n$是X的样本，$\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$.

$$\mu_1=E(X)=\frac{a+b}{2},$$

$$\mu_2=E(X^2)=D(X)+[E(X)]^2=\frac{(b-a)^2}{12}+\frac{(a+b)^2}{4},$$

令 $\mu_1=A_1=\overline{X}$，$\mu_2=A_2=\frac{1}{n}\sum_{i=1}^{n}X_i^2$，则有

$$\begin{cases}a+b=2A_1,\\ b-a=\sqrt{12(A_2+A_1^2)},\end{cases}$$

解方程组，得

$$\hat{a}=A_1-\sqrt{3(A_2-A_1^2)}=\overline{X}-\sqrt{\frac{3}{n}\sum_{i=1}^{n}(X_i-\overline{X})^2},$$

$$\hat{b}=A_1+\sqrt{3(A_2-A_1^2)}=\overline{X}+\sqrt{\frac{3}{n}\sum_{i=1}^{n}(X_i-\overline{X})^2}.$$

例4 设总体$X\sim B(N,p)$，N与p为未知参数，$X_1,X_2,\cdots,X_n$为X的一个样本，求N,p的矩估计量.

解 因为$E(X)=Np$，$D(X)=Np(1-p)$，所以

$$E(X^2)=D(X)+[E(X)]^2=Np(1-p)+N^2p^2.$$

令 $\mu_1=A_1$，$\mu_2=A_2$，得方程组

$$\begin{cases}\overline{X}=Np,\\ \frac{1}{n}\sum_{i=1}^{n}X_i^2=N^2p^2+Np(1-p).\end{cases}$$

解方程组，得

$$\hat{p}=1+\overline{X}-\frac{1}{n\overline{X}}\sum_{i=1}^{n}X_i^2=1-\frac{B_2}{\overline{X}}.$$

$$\hat{N}=\frac{\overline{X}}{\hat{p}}=\frac{\overline{X}^2}{\overline{X}-B_2}\quad\left(B_2=\frac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})^2\right).$$

例 5 设使用了某种仪器对同一量进行了 12 次独立的测量，其数据(单位:mm)如下：

232.50，232.48，232.15，232.53，232.45，232.30，

232.48，232.05，232.45，232.60，232.47，232.30，

试用矩估计法估计测量值的真值与方差(设仪器无系统误差).

解 因为仪器无系统误差，所以

$$\hat{\theta}=\hat{\mu}=\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i=232+\frac{1}{12}\sum_{i=1}^{n}(X_i-232)$$

$$=232+\frac{1}{12}\times 4.76=232.3967.$$

用样本二阶中心矩 B_2 估计方差 σ^2，有

$$\hat{\sigma}^2=\frac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})^2=\frac{1}{n}\sum_{i=1}^{n}(X_i-a)^2-(\overline{X}-a)^2$$

$$=\frac{1}{12}\sum_{i=1}^{12}(X_i-232)^2-(232.3967-232)^2$$

$$=0.1819-0.1574=0.0245.$$

例 6 设 $X_1,X_2,\cdots,X_n$ 为总体 X 的一个样本，求 X 的概率密度为下述情形时参数的矩估计量：

(1) $f(x)=\begin{cases}\theta c^{\theta}x^{-(\theta+1)}, & x>c,\\ 0, & \text{其它}\end{cases}$ (θ 未知，c 已知)；

(2) $f(x)=\begin{cases}\frac{1}{\theta}\mathrm{e}^{-(x-\mu)/\theta}, & x\geqslant\mu,\\ 0, & \text{其它}\end{cases}$ (θ,μ 均未知).

解 (1) $E(X)=\int_0^{+\infty}\theta c^{\theta}x^{-(\theta+1)}x\mathrm{d}x=c\,\frac{\theta}{1-\theta}$,

故，令 $E(X)=\overline{X}$，得 $\hat{\theta}=\dfrac{\overline{X}}{\overline{X}-c}$.

(2) $E(X)=\int_{\mu}^{+\infty}\dfrac{1}{\theta}x\mathrm{e}^{-(x-\mu)/\theta}\mathrm{d}x=\mu+\theta$,

故，令 $E(X)=\overline{X}$，得 $\hat{\mu}=\overline{X}-\mu$.

$$E(X^2)=\int_{\mu}^{+\infty}\frac{1}{\theta}x^2\mathrm{e}^{-(x-\mu)/\theta}\mathrm{d}x=\mu^2+2\theta(\mu+\theta),$$

故，令 $E(X^2)=B^2+A_1^2$，可解得

$$\hat{\theta}=\sqrt{\frac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})^2},\quad \hat{\mu}=\overline{X}-\sqrt{\frac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})^2}.$$

例 7 设总体 X 服从几何分布 $G(p)$，分布律为 $P\{X=x\}=(1-p)^{x-1}p$，$x=1,2,\cdots$，其中 p 为未知参数 $(0<p<1)$，求 p 的矩估计量.

解 因为
$$\overline{X}=E(X)=\sum_{i=1}^{\infty}x(1-p)^{x-1}p$$
$$=\sum_{i=1}^{\infty}(kq^{k-1})p=p\,\frac{1}{p^2}=\frac{1}{p},$$

所以
$$\hat{p}=1/\overline{X}.$$

例 8 设总体 X 的概率密度为

$$f(x)=\begin{cases}\dfrac{\beta^k}{(k-1)!}x^{k-1}\mathrm{e}^{-\beta x}, & x>0,\\ 0, & x\leqslant 0,\end{cases}$$

其中 k 为已知整数，β 为未知参数，求 β 的矩估计量.

解 $E(X)=\int_{0}^{+\infty}x\,\dfrac{\beta^k}{(k-1)!}x^{k-1}\mathrm{e}^{-\beta x}\mathrm{d}x$ （令 $\beta x=y$）

$$=\frac{1}{\beta(k-1)!}\int_{0}^{+\infty}y^k\mathrm{e}^{-y}\mathrm{d}y=\frac{\Gamma(k+1)}{\beta(k-1)!}$$
$$=\frac{k!}{\beta(k-1)!}=\frac{k}{\beta},$$

令 $E(X)=\overline{X}$，得 β 的矩估计量 $\hat{\beta}=k/\overline{X}$.

例 9 设总体 X 的概率密度为

$$f(x)=\frac{1}{2\lambda}e^{-|x-\theta|/\lambda},\ \lambda>0,$$

求θ和λ的矩估计量.

解
$$E(X)=\int_{-\infty}^{+\infty}\frac{1}{2\lambda}e^{-|x-\theta|/\lambda}x\mathrm{d}x=\theta,$$

$$D(X)=\int_{-\infty}^{+\infty}\frac{(x-\theta)^2}{2\lambda}e^{-|x-\theta|/\lambda}\mathrm{d}x=\frac{\lambda^2}{2}\int_{-\infty}^{+\infty}u^2e^{-|u|}\mathrm{d}u$$

$$=\lambda^2\int_{-\infty}^{+\infty}u^2e^{-u}\mathrm{d}u=\lambda^2\Gamma(3)=2\lambda^2.$$

令
$$\overline{X}=E(X),\quad B_2=\frac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})^2=D(X),$$

则有
$$\begin{cases}\hat{\theta}=\overline{X},\\ B_2=2\hat{\lambda}^2,\end{cases}$$

解得
$$\hat{\theta}=\overline{X},\quad \hat{\lambda}=\sqrt{\frac{1}{2n}\sum_{i=1}^{n}(X_i-\overline{X})^2}.$$

二、极大似然估计的求法

在建立了样本的似然函数$L(\theta)$后，在$p\{x;\theta\}$或$f\{x;\theta\}$可微时，可以利用微积分学中求极值的求法，令

$$\frac{\mathrm{d}}{\mathrm{d}\theta}L(\theta)=0\quad 或\quad \frac{\mathrm{d}}{\mathrm{d}\theta}\ln L(\theta)=0$$

($L(\theta)$与$\ln L(\theta)$有相同的极值点)，即可解得$\hat{\theta}$. 若有多个未知参数$\theta_1,\theta_2,\cdots,\theta_k$，则令

$$\frac{\partial}{\partial\theta_i}L=0\quad 或\quad \frac{\partial}{\partial\theta_i}\ln L=0\quad (i=1,2,\cdots,k),$$

解得$\hat{\theta}_1,\hat{\theta}_2,\cdots,\hat{\theta}_k$.

从理论上说，求得的解只满足极值的必要条件，要证明是极大，则还应验证是否满足极值的充分条件. 但一般不考虑此项.

在$p\{x;\theta\}$或$f\{x;\theta\}$不可微时，只能由定义来确定极大似然估计量.

极大似然估计和矩估计有一条重要性质，即：若$\hat{\theta}$是θ的极大

似然估计或矩估计，则 $g(\hat{\theta})$ 也是 $g(\theta)$ 的极大似然估计或矩估计.

例 10 设总体 X 的概率分布为

X	1	2	3
p_k	θ^2	$2\theta(1-\theta)$	$(1-\theta)^2$

其中 θ 为未知参数. 现抽得一个样本 $x_1=1, x_2=2, x_3=1$，求 θ 的极大似然估计值.

解 建立样本的似然函数

$$L(\theta)=\prod_{i=1}^{3} p(x_i;\theta)=\theta^2\cdot 2(1-\theta)\theta\cdot\theta^2=2\theta^5(1-\theta),$$

取对数，得 $$\ln L(\theta)=\ln 2+5\ln\theta+\ln(1-\theta),$$

求导数，得 $$\frac{\mathrm{d}}{\mathrm{d}\theta}\ln L(\theta)=\frac{5}{\theta}-\frac{1}{1-\theta}.$$

令上式等于零，所以 θ 的极大似然估计值为 $\hat{\theta}=5/6$.

例 11 设总体 X 服从几何分布 $G(p)$，分布律为

$$P\{X=x\}=(1-p)^{x-1}p,\ x=1,2,\cdots,$$

其中 $p\ (0<p<1)$ 为未知参数，求 p 的极大似然估计量.

解 建立样本的似然函数

$$L(p)=\prod_{i=1}^{n}\left[p(1-p)^{x_i-1}\right]=p^n(1-p)^{\sum_{i=1}^{n}x_i-n},$$

取对数，得 $$\ln L(p)=n\ln p+\left(\sum_{i=1}^{n}x_i-n\right)\ln(1-p),$$

求导数，得 $$\frac{\mathrm{d}}{\mathrm{d}p}\ln L(p)=\frac{n}{p}-\frac{\sum_{i=1}^{n}x_i-n}{1-p}.$$

令上式等于零，所以 p 的极大似然估计量 $\hat{p}=1/\overline{X}$.

例 12 设 $X\sim U(a,b)$，a,b 为未知，$x_1,x_2,\cdots,x_n$ 是一组样本观察值，求 a,b 的极大似然估计值与极大似然估计量.

解 因为 $X\sim U(a,b)$，所以

$$f(x;a,b)=\begin{cases}1/(b-a), & a\leqslant x\leqslant b,\\ 0, & \text{其它}.\end{cases}$$

取 $x_{(1)}=\min(x_1,x_2,\cdots,x_n)$, $x_{(n)}=\max(x_1,x_2,\cdots,x_n)$,

则有 $a\leqslant x_{(1)},x_{(n)}\leqslant b$. 于是,似然函数

$$L(a,b)=\begin{cases}1/(b-a)^n, & a\leqslant x_{(1)},x_{(n)}\leqslant b,\\ 0, & \text{其它},\end{cases}$$

则对满足条件 $a\leqslant x_{(1)},x_{(n)}\leqslant b$ 的任意 a,b,有

$$L(a,b)=1/(b-a)^n\leqslant 1/[x_{(n)}-x_{(1)}]^n.$$

即 $L(a,b)$ 在 $a=x_{(1)},b=x_{(n)}$ 时取得极大值 $[x_{(n)}-x_{(1)}]^{-n}$,所以 a,b 的极大似然估计值为

$$\hat{a}=x_{(1)}=\min_{1\leqslant i\leqslant n}x_i,\quad \hat{b}=x_{(n)}=\max_{1\leqslant i\leqslant n}x_i.$$

a,b 的极大似然估计量为

$$\hat{a}=\min_{1\leqslant i\leqslant n}X_i,\quad \hat{b}=\max_{1\leqslant i\leqslant n}X_i.$$

例13 设 $X_1,X_2,\cdots,X_n$ 为来自总体 X 的一个样本,X 的概率密度为

$$f(x;\theta,\mu)=\begin{cases}\dfrac{1}{\theta}e^{-(x-\mu)/\theta}, & x\geqslant\mu,\\ 0, & \text{其它}\end{cases}\quad(\theta>0),$$

求未知参数 θ 和 μ 的极大似然估计量.

解 建立样本的似然函数

$$L(x;\mu,\theta)=\theta^{-n}e^{-\sum\limits_{i=1}^{n}(x_i-\mu)/\theta},$$

当 $x_1,x_2,\cdots,x_n$ 取定时,μ 的最大值在

$$\hat{\mu}=x_{(1)}=\min(x_1,x_2,\cdots,x_n)$$

时取得,此时 $L(\hat{\mu},\theta)=\max\limits_{\mu}L(\mu,\theta)$.

代入似然函数并取对数,得

$$\ln L(\hat{\mu},\theta)=-n\ln\theta-\frac{1}{\theta}\sum_{i=1}^{n}(x_i-x_{(1)}),$$

求偏导数,得 $$\frac{\partial\ln L}{\partial\theta}=-\frac{n}{\theta}+\frac{1}{\theta^2}\sum_{i=1}^{n}(x_i-x_{(1)}),$$

令上式等于零,解得

$$\hat{\theta}=\frac{1}{n}\sum_{i=1}^{n}(x_i-x_{(1)})=\bar{x}-x_{(1)}.$$

即 μ,θ 的极大似然估计量分别为 $\hat{\mu}=X_{(1)},\hat{\theta}=\overline{X}-X_{(1)}$.

例14 设在n 次独立试验中事件A 发生了k 次，求事件A 发生的概率 p 的极大似然估计量.

解 设$x_1,x_2,\cdots,x_n$ 为 X 的一组样本观察值，$X\sim B(1,p)$，其分布律为$P\{X=x\}=p^{x_i}(1-p)^{1-x}$，$x=1,2$. 建立似然函数

$$L(p)=\prod_{i=1}^{n}p^{x_i}(1-p)^{1-x_i}=p^{\sum\limits_{i=1}^{n}x_i}(1-p)^{n-\sum\limits_{i=1}^{n}x_i},$$

取对数，得 $\ln L(p)=\sum\limits_{i=1}^{n}x_i\ln p+\left(n-\sum\limits_{i=1}^{n}x_i\right)\ln(1-p)$，

求导数，得 $\frac{\mathrm{d}}{\mathrm{d}p}\ln L(p)=\frac{1}{p}\sum\limits_{i=1}^{n}x_i-\frac{1}{1-p}\left(n-\sum\limits_{i=1}^{n}x_i\right)$，

令上式等于零，解得 p 的极大似然估计值为 $\hat{p}=\frac{1}{n}\sum\limits_{i=1}^{n}x_i=\bar{x}$. p 的极大似然估计量为

$$\hat{p}=\frac{1}{n}\sum_{i=1}^{n}X_i=\overline{X}.$$

例15 设$X\sim N(\mu,\sigma^2)$，μ,σ^2 为未知参数，$x_1,x_2,\cdots,x_n$ 是来自X 的一个样本值，求 μ,σ^2 的极大似然估计值和极大似然估计量.

解 因为$f(x)=\frac{1}{\sqrt{2\pi}\sigma}\mathrm{e}^{-(x-\mu)^2/(2\sigma^2)}$，$-\infty<x<+\infty$，所以，其似然函数为

$$L(\mu,\sigma)=\prod_{i=1}^{n}\frac{1}{\sqrt{2\pi}\sigma}\mathrm{e}^{-(x_i-\mu)^2/(2\sigma^2)}=(2\pi\sigma^2)^{-n/2}\mathrm{e}^{-\sum\limits_{i=1}^{n}(x_i-\mu)^2/(2\sigma^2)},$$

取对数，得

$$\ln L=-\frac{n}{2}\ln 2\pi-\frac{n}{2}\ln\sigma^2-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2,$$

求偏导数，得 $$\begin{cases}\dfrac{\partial}{\partial\mu}\ln L=\dfrac{1}{\sigma^2}\left(\sum\limits_{i=1}^{n}x_i-n\mu\right),\\ \dfrac{\partial}{\partial\sigma^2}\ln L=-\dfrac{n}{2\sigma^2}+\dfrac{1}{2\sigma^4}\sum\limits_{i=1}^{n}(x_i-\mu)^2,\end{cases}$$

令上两式分别等于零，解得极大似然估计值为

$$\hat{\mu}=\bar{x},\quad \hat{\sigma}^2=\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2,$$

所以极大似然估计量为

$$\hat{\mu}=\overline{X},\quad \hat{\sigma}^2=\frac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})^2.$$

例 16 设随机变量 X 的概率密度为

$$f(x;\mu,\sigma^2)=\frac{1}{\sqrt{2\pi}\sigma x}e^{-(\ln x-\mu)^2/(2\sigma^2)},\ x>0,$$

求未知参数 μ 和 σ^2 的极大似然估计量和矩估计量.

解 (1) 建立样本的似然函数

$$L(\mu,\sigma^2)=(2\pi\sigma^2)^{-n/2}\frac{1}{X_1X_2\cdots X_n}e^{-\sum\limits_{i=1}^{n}(X_i-\mu)^2/(2\sigma^2)},$$

取对数，得

$$\ln L=\frac{-n}{2}\ln 2\pi-\frac{n}{2}\ln\sigma^2-\sum_{i=1}^{n}\ln X_i-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(\ln X_i-\mu)^2,$$

求偏导数，得

$$\begin{cases}\dfrac{\partial}{\partial\mu}\ln L=\dfrac{1}{\sigma^2}\sum\limits_{i=1}^{n}(\ln X_i-\mu),\\ \dfrac{\partial}{\partial\sigma^2}\ln L=-\dfrac{n}{2}\dfrac{1}{\sigma^2}+\dfrac{1}{2\sigma^4}\sum\limits_{i=1}^{n}(\ln X_i-\mu)^2,\end{cases}$$

令上两式分别等于零，其解为 μ 和 σ^2 的极大似然估计量，即

$$\hat{\mu}=\frac{1}{n}\sum_{i=1}^{n}\ln X_i=\overline{\ln X},\quad \hat{\sigma}^2=\frac{1}{n}\sum_{i=1}^{n}(\ln X_i-\overline{\ln X})^2.$$

(2) 因为 $E(X)=e^{\mu+\sigma^2/2}$，$D(X)=e^{\sigma^2+2\mu}(e^{\sigma^2}-1)$(见第四章第一节例 24). 令 $\overline{X}=E(X)$，$B_2=D(X)$，得

$$\begin{cases}\overline{X}=e^{\hat{\mu}+\hat{\sigma}^2/2},\\ B_2=e^{2\hat{\mu}+\hat{\sigma}^2}(e^{\hat{\sigma}^2}-1)=\overline{X}^2(e^{\hat{\sigma}^2}-1),\end{cases}$$

解得 σ^2 的矩估计量为

$$\hat{\sigma}^2=\ln(1+B_2/\overline{X}^2),$$

μ 的矩估计量为

$$\hat{\mu}=\ln[\overline{X}^2/\sqrt{B_2+\overline{X}^2}]\quad \left(B_2=\frac{1}{n}\sum_{i=1}^{2}(X_i-\overline{X})^2\right).$$

例 17 设某种灯泡寿命服从正态分布.在某天生产的灯泡中抽取10只,测得寿命(单位:h)为

1067, 919, 1196, 785, 1126, 936, 918, 1156, 920, 948,

若总体 X 参数未知,试用极大似然估计法求该天生产的灯泡能使用1200 h以上的概率.

解 总体 $X\sim N(\mu,\sigma^2)$,μ,σ^2 为未知参数,由例15知

$$\hat{\mu}=\overline{X},\quad \hat{\sigma}^2=\frac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})^2,$$

则
$$\begin{aligned}P\{X>1200\}&=1-P\{X\leqslant 1200\}=1-\Phi\left(\frac{1200-\hat{\mu}}{\hat{\sigma}}\right)\\&=1-\Phi\left(\frac{1200-997.1}{124.8}\right)=1-\Phi(1.625)\\&=1-0.978=0.022,\end{aligned}$$

所以,灯泡能使用1200 h以上的概率为0.022.

例 18 为了估计湖中鱼的条数 N,先从湖中捕捉 r 条鱼,做上记号后放回湖中.过一段时间后,再从湖中捕出 s $(s>r)$ 条,发现其中有 t $(0\leqslant t\leqslant r)$ 条标有记号.试以此估计湖中鱼的条数 N 的值.

解 为了启发读者思维,我们给出下面几种解法.

(1) 矩估计法

$E(X)=rs/N$(见第四章第一节例5).在只捕一次的情况下,捕到了 t 条有记号的鱼,将它看作一个样本观察值.令总体一阶矩等于样本一阶矩(原点矩),则 $rs/N=t$,于是得 N 的矩估计量为

$$\hat{N}=[rs/t].$$

（2）极大似然估计法

设在捕到的s条鱼中，有记号的鱼数是一个取值为$0,1,2,\cdots,r$的随机变量X.

考虑在s条鱼中恰有t条鱼有记号的概率，因为X服从超几何分布，分布律为

$$P\{X=k\}=C_r^kC_{N-r}^{s-k}/C_N^s,\ k=0,1,2,\cdots,r,$$

故
$$P\{X=t\}=C_r^tC_{N-r}^{s-t}/C_N^s \xlongequal{\text{令}} L(N).$$

这里N是未知参数. 按极大似然估计法，要找到$\hat{N}$，使$L(N)$为最大. 为此，考虑比值

$$\begin{aligned}R(N)&=\frac{L(N)}{L(N-1)}=\frac{C_r^tC_{N-r}^{s-t}}{C_N^s}\Bigg/\frac{C_r^tC_{N-r-1}^{s-t}}{C_{N-1}^s}\\&=\frac{C_{N-r}^{s-t}C_{N-1}^s}{C_N^sC_{N-r-1}^{s-t}}=\frac{(N-r)(N-s)}{N(N-r-s+1)}\\&=\frac{N^2-Nr-Ns+rs}{N^2-Nr-Ns+Nt},\end{aligned}$$

显然，当$rs<Nt$时，$R(N)<1$；当$rs>Nt$时，$R(N)>1$. 即$L(N)$在N经过rs/t时，由增加转为减少，亦即，当$N=rs/t$时，$L(N)$取得极大值. 所以，N的极大似然估计量为$\hat{N}=[rs/t]$.

（3）比例法（用频率估计）

依题意，湖中有记号的鱼所占比例应为r/N，而在捕到的s条鱼中，有记号的鱼的频率（比例）是t/s. 由于捕鱼是随机的，每条鱼是独立的，可以认为，应该有

$$r/N=t/s\Longrightarrow N=sr/t.$$

从而，N的估计量为$\hat{N}=[rs/t]$.

三、估计量的评选

估计量的评选，即讨论估计量是否满足估计量的三个标准，所以，一般可以直接依无偏性、有效性、一致性的定义来确定. 但无偏性是数学期望，有效性是方差，所以又可以利用数学期望与方差的

运算性质来评选. 读者要注意灵活运用.

例 19 设总体 $X\sim N(\mu,\sigma^2)$，X_1,X_2,X_3 是 X 的一个样本，又

$$\hat{\mu}_1=\frac{1}{5}X_1+\frac{3}{10}X_2+\frac{1}{2}X_3,\quad \hat{\mu}_2=\frac{1}{3}X_1+\frac{1}{4}X_2+\frac{5}{12}X_3,$$

$$\hat{\mu}_3=\frac{1}{3}X_1+\frac{1}{6}X_2+\frac{1}{2}X_3,$$

验证 $\hat{\mu}_1,\hat{\mu}_2,\hat{\mu}_3$ 都是 μ 的无偏估计量，并判断哪个最有效.

解 因为 X_1,X_2,X_3 相互独立且同分布，有

$$E(X_1)=E(X_2)=E(X_3)=\mu,$$

所以

$$E(\hat{\mu}_1)=\frac{1}{5}\mu+\frac{3}{10}\mu+\frac{1}{2}\mu=\mu,$$

$$E(\hat{\mu}_2)=\frac{1}{3}\mu+\frac{1}{4}\mu+\frac{5}{12}\mu=\mu,$$

$$E(\hat{\mu}_3)=\frac{1}{3}\mu+\frac{1}{6}\mu+\frac{1}{2}\mu=\mu,$$

知 $\hat{\mu}_1,\hat{\mu}_2,\hat{\mu}_3$ 都是 μ 的无偏估计量. 又

$$D(\hat{\mu}_1)=\frac{1}{25}\sigma^2+\frac{9}{100}\sigma^2+\frac{1}{4}\sigma^2=\frac{684}{1800}\sigma^2,$$

$$D(\hat{\mu}_2)=\frac{1}{9}\sigma^2+\frac{1}{16}\sigma^2+\frac{25}{144}\sigma^2=\frac{625}{1800}\sigma^2,$$

$$D(\hat{\mu}_3)=\frac{1}{9}\sigma^2+\frac{1}{36}\sigma^2+\frac{1}{4}\sigma^2=\frac{700}{1800}\sigma^2,$$

故 $D(\hat{\mu}_2)<D(\hat{\mu}_1)<D(\hat{\mu}_3)$，$\hat{\mu}_2$ 是最有效估计量.

例 20 设总体 $X\sim N(\mu,\sigma^2)$，$X_1,X_2,\cdots,X_n$ 是 X 的一个样本，试确定常数 C，使 $C\sum\limits_{i=1}^{n-1}(X_{i+1}-X_i)^2$ 为 σ^2 的无偏估计.

解 依题意，即要使 $E\left[C\sum\limits_{i=1}^{n-1}(X_i-X_i)^2\right]=\sigma^2$，但

$$E\left[C\sum_{i=1}^{n-1}(X_i-X_i)^2\right]=C\sum_{i=1}^{n-1}[E(X_{i+1}^2)-2E(X_{i+1})E(X_i)+E(X_i^2)]$$

$$=2C\sum_{i=1}^{n-1}\{E(X^2)-[E(X)]^2\}$$

$$=2C\sum_{i=1}^{n-1}D(X)=2C(n-1)\sigma^2,$$

故 $C=\frac{1}{2(n-1)}$. 此时 $\frac{1}{2n-1}\sum_{i=1}^{n-1}(X_{i+1}-X_i)^2$ 是 σ^2 的无偏估计量.

例 21 设 $X_1,X_2,\cdots,X_n$ 是总体 X 的一个样本,验证估计量 $\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$ 和 $W=\sum_{i=1}^{n}a_iX_i$ $\left(\sum_{i=1}^{n}a_i=1\right)$ 都是 $E(X)$ 的无偏估计量,且 $\overline{X}$ 比 W 有效.

解 $E(\overline{X})=\frac{1}{n}\sum_{i=1}^{n}E(X_i)=\frac{1}{n}nE(X)=E(X)$,

$$E(W)=\sum_{i=1}^{n}a_iE(X_i)=E(X)\sum_{i=1}^{n}a_i=E(X),$$

可知,$\overline{X}$ 与 W 都是 $E(X)$ 的无偏估计量. 而

$$D(\overline{X})=\frac{1}{n^2}\sum_{i=1}^{n}D(X_i)=\frac{1}{n^2}nD(X)=\frac{1}{n}D(X),$$

$$D(W)=\sum_{i=1}^{n}a_i^2D(X_i)=D(X)\sum_{i=1}^{n}a_i^2\geqslant\frac{1}{2}D(X),$$

因为

$$(n-1)\sum_{i=1}^{n}a_i^2\geqslant 2\sum_{1\leqslant i<j\leqslant n}a_ia_j\Longrightarrow\sum_{i=1}^{n}a_i^2\geqslant\frac{1}{n}\left(\sum_{i=1}^{n}a_i^2+2\sum_{1\leqslant i<j\leqslant n}a_ia_j\right),$$

所以,$\overline{X}$ 比 W 更有效.

例22 设 $X_1,X_2,\cdots,X_n$ 是总体 $X\sim U(0,\theta)$ 的一个样本,证明:

(1) $\hat{\theta}_1=2\overline{X}$ 与 $\hat{\theta}_2=\frac{n+1}{n}X_{(n)}$ 是 θ 的无偏估计;

(2) $\hat{\theta}_2$ 比 $\hat{\theta}_1$ 更有效($n\geqslant 2$).

证 (1) 因为 $X\sim U(0,\theta)$,所以

$$f(x)=\begin{cases}1/\theta, & 0<x<\theta,\\ 0, & \text{其它}.\end{cases}$$

知

$$f_{X(n)}(x)=\begin{cases}nx^{n-1}/\theta^n, & 0<x<\theta,\\ 0, & \text{其它},\end{cases}$$

所以 $$E(2\overline{X})=\frac{2}{n}\sum_{i=1}^{n}X_i=\frac{2}{n}n\cdot\frac{\theta+0}{2}=\theta,$$

$$E\left(\frac{n+1}{n}X_{(n)}\right)=\frac{n+1}{n}\int_0^\theta\frac{nx^n}{\theta^n}\mathrm{d}x=\frac{n+1}{n}\cdot\frac{n}{n+1}\theta=\theta.$$

从而知 $\hat{\theta}_1$ 和 $\hat{\theta}_2$ 都是 θ 的无偏估计量.

(2) 因为 $D(\hat{\theta}_1)=D(2\overline{X})=\frac{4}{n}\cdot\frac{(\theta-0)^2}{12}=\frac{\theta^2}{3n}$,

$$D\left(\frac{n+1}{n}X_{(n)}\right)=\frac{(n+1)^2}{n^2}\{E(X_{(n)}^2)-[E(X_{(n)})]^2\},$$

而 $$E(X_{(n)}^2)=\int_0^\theta\frac{nx^{n+1}}{\theta^n}\mathrm{d}x=\frac{n}{n+2}\theta^2,$$

所以

$$D\left(\frac{n+1}{n}X_{(n)}\right)=\frac{(n+1)^2}{n^2}\left[\frac{n}{n+2}\theta^2-\left(\frac{n}{n+1}\theta\right)^2\right]=\frac{\theta^2}{n(n+2)}.$$

当 $n\geqslant 2$ 时,$D(\hat{\theta}_1)>D(\hat{\theta}_2)$,所以 $\hat{\theta}_2$ 比 $\hat{\theta}_1$ 更有效.

例 23 设总体 $X\sim\pi(\lambda)$,$\lambda>0$,$X_1,X_2,\cdots,X_n$ 是 X 的一个样本,证明:

(1) $\overline{X}$ 是 λ 的无偏估计量,但 $(\overline{X})^2$ 不是 λ^2 的无偏估计量;

(2) $\overline{X}$ 是 λ 的达到方差界的无偏估计.

证 (1) 因为 $E(X_i)=E(X)=\lambda$,$D(X_i)=D(X)=\lambda$,所以

$$E(\overline{X})=\frac{1}{n}\sum_{i=1}^{n}E(X_i)=\frac{1}{n}n\lambda=\lambda,$$

故 $\overline{X}$ 是 λ 的无偏估计.

$$D(\overline{X})=\frac{1}{n^2}\sum_{i=1}^{n}D(X_i)=\frac{1}{n^2}n\lambda=\frac{\lambda}{n},$$

$$E(\overline{X}^2)=D(X)+[E(\overline{X})]^2=\frac{\lambda}{n}+\lambda^2\neq\lambda,$$

故 $\overline{X}^2$ 不是 λ^2 的无偏估计量.

一般地,如果 $\hat{\theta}$ 是 θ 的极大似然估计或矩估计,则 $g(\hat{\theta})$ 也是 $g(\theta)$ 的极大似然估计或矩估计. 但这一性质对无偏估计不成立.

(2) 因为

$$p(x;\lambda)=\frac{\lambda^x}{x!}e^{-\lambda},$$

$$\ln p(x;\lambda)=x\ln\lambda-\lambda-\ln(x!),\quad \frac{\partial}{\partial\lambda}\ln p(x;\lambda)=\frac{x}{\lambda}-1,$$

而 $$E\left[\left(\frac{x}{\lambda}-1\right)^2\right]=\sum_{x=0}^{\infty}\left(\frac{x}{\lambda}-1\right)^2\frac{\lambda^x}{x!}e^{-\lambda}$$

$$=\sum_{x=0}^{\infty}\frac{x^2}{\lambda^2}\frac{\lambda^x}{x!}e^{-\lambda}-\sum_{x=0}^{\infty}\frac{2x}{\lambda}\frac{\lambda^x}{x!}e^{-\lambda}+\sum_{x=0}^{\infty}\frac{\lambda^x}{x!}e^{-\lambda},$$

由 $$E(X)=\lambda,\quad E(X^2)=\lambda+\lambda^2,\quad D(X)=\lambda,$$

知 $$E\left[\left(\frac{x}{\lambda}-1\right)^2\right]=\frac{\lambda+\lambda^2}{\lambda^2}-\frac{2\lambda}{\lambda}+1=\frac{1}{\lambda},$$

而 $$D(\overline{X})=\frac{\lambda}{n}=\frac{1}{nE\left\{\left[\frac{\partial}{\partial\lambda}\ln p(x;\lambda)\right]^2\right\}}=\frac{1}{n\cdot\frac{1}{\lambda}}=\frac{\lambda}{n},$$

所以，$\overline{X}$ 是 λ 的达到方差界的无偏估计.

例 24 设总体 $X\sim B(1,p)$，其中 p 为未知参数，$0<p<1$，X_1，X_2，…，X_n 是 X 的一个样本，证明 $\overline{X}$ 是 p 的有效估计量.

证 因为 $P(x;p)=p^x(1-p)^{1-x}$，$x=0,1$，所以

$$\frac{\partial}{\partial p}\ln P(x;p)=\frac{\partial}{\partial p}[x\ln p+(1-x)\ln(1-p)]$$

$$=\frac{x}{p}+\frac{1-x}{1-p}=\frac{x-p}{p(1-p)}.$$

而 $$E\left[\frac{x-p}{p(1-p)}\right]^2=\frac{1}{p^2(1-p)^2}E[(X-p)^2]=\frac{1}{p(1-p)}.$$

又 $$\hat{p}=\overline{X},\quad D(\hat{p})=D(\overline{X})=p(1-p)=1\Big/E\left[\frac{x-p}{p(1-p)}\right]^2,$$

由克莱姆-拉奥不等式

$$D(\hat{\theta})=1\Big/\left\{nE\left[\left(\frac{\partial\ln p(x;\theta)}{\partial\theta}\right)^2\right]\right\}$$

可知，$\overline{X}$ 是 p 的达到方差界的无偏估计，所以是 p 的有效估计量.

例 25 设 $X_1,X_2,\cdots,X_{n_1}$ 为 $X\sim N(\mu_1,\sigma^2)$ 的一个样本，Y_1,Y_2，…，Y_{n_2} 是 $Y\sim N(\mu_2,\sigma^2)$ 的一个样本，且相互独立，S_1^2,S_2^2 分别为它

们的样本方差. 证明：对于任意常数 $a,b(a+b=1)$，$Z=aS_1^2+bS_2^2$ 都是 σ^2 的无偏估计，并确定 a,b 的值，使 $D(Z)$ 达到最小.

证 因为

$$E(S_1^2)=\sigma^2,\quad E(S_2^2)=\sigma^2,$$

$$\frac{(n_1-1)}{\sigma^2}S_1^2\sim\chi^2(n_1-1),\quad \frac{(n_2-1)}{\sigma^2}S_2^2\sim\chi^2(n_2-1),$$

且相互独立，所以

$$D(S_1^2)=\frac{2\sigma^4}{(n_1-1)},\quad D(S_2^2)=\frac{2\sigma^4}{(n_2-1)}.$$

在 $a+b=1$ 时，$E(Z)=aE(S_1^2)+bE(S_2^2)=\sigma^2$，故 Z 是 σ^2 的无偏估计.

$$D(Z)=D(aS_1^2+bS_2^2)=\left(\frac{a^2}{n_1-1}+\frac{b^2}{n_2-1}\right)2\sigma^4$$

$$=\left[\frac{a^2}{n_1-1}+\frac{(1-a)^2}{n_2-1}\right]2\sigma^4,$$

$$\frac{\mathrm{d}D(Z)}{\mathrm{d}a}=2\sigma^4\left[\frac{2a}{n_1-1}-\frac{2(1-a)}{n_2-1}\right].$$

令上式等于零，并由 a 与 b 的对称性，得

$$a=\frac{n_1-1}{n_1+n_2-2},\quad b=\frac{n_2-1}{n_1+n_2-2}.$$

又由
$$\frac{\mathrm{d}^2D(Z)}{\mathrm{d}a^2}=2\sigma^4\left(\frac{2}{n_1-1}+\frac{2}{n_2-1}\right)>0$$
知，所确定的 a,b 值能使 $D(Z)$ 达到最小，此时

$$Z=\frac{1}{n_1+n_2-2}[(n_1-1)S_1^2+(n_2-1)S_2^2]$$

具有最小方差.

例 26 设总体 X 的一阶矩和二阶矩存在，分布是任意的. 记 $E(X)=\mu,D(X)=\sigma^2$，样本均值 $\overline{X}$ 与 $\frac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})^2$ 是 μ 与 σ^2 的矩估计量，问：它们是否是 μ 与 σ^2 的无偏估计量？

解 因为

$$E(\overline{X})=E\left[\frac{1}{n}\sum_{i=1}^{n}X_i\right]=\frac{1}{n}\sum_{i=1}^{n}E(X_i)=\frac{1}{n}n\mu=\mu,$$

所以 $\overline{X}$ 是 μ 的无偏估计量.

$$E\left[\frac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})^2\right]$$

$$=\frac{1}{n}E\left\{\sum_{i=1}^{n}[(X_i-\mu)-(\overline{X}-\mu)]^2\right\}$$

$$=\frac{1}{n}E\left[\sum_{i=1}^{n}(X_i-\mu)^2-2\sum_{i=1}^{n}(X_i-\mu)(\overline{X}-\mu)+n(\overline{X}-\mu)^2\right]$$

$$=\frac{1}{n}E\left[\sum_{i=1}^{n}(X_i-\mu)^2-n(\overline{X}-\mu)^2\right]$$

$$=\frac{1}{n}\left[\sum_{i=1}^{n}D(X_i)-nD(\overline{X})\right]=\frac{1}{n}(n\sigma^2-\sigma^2)=\frac{n-1}{n}\sigma^2,$$

所以 $\frac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})^2$ 不是 σ^2 的无偏估计量,由于

$$\lim_{n\to\infty}\frac{n-1}{n}\sigma^2=\sigma^2,$$

故 $\hat{\sigma}^2=\frac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})^2$ 是 σ^2 的渐近无偏估计量.

但是,对样本方差 $S^2=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2$,有

$$E(S^2)=E\left[\frac{n}{n-1}\cdot\frac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})^2\right]$$

$$=\frac{n}{n-1}E\left[\frac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})^2\right]=\frac{n}{n-1}\cdot\frac{n-1}{n}\sigma^2=\sigma^2,$$

所以,样本方差是 σ^2 的无偏估计量.

例 27 设 $X_1,X_2,\cdots,X_{n_1}$ 是 $X\sim N(\mu_1,\sigma^2)$ 的一个样本,$Y_1,Y_2,\cdots,Y_{n_2}$ 是 $Y\sim N(\mu_2,\sigma^2)$ 的一个样本,两样本相互独立,μ_1,μ_2 为未知参数.

(1) 求参数 $\mu_1-\mu_2$ 的一个无偏估计;

（2）证明：$S_W^2=\frac{1}{n_1+n_2-2}\left[\sum_{i=1}^{n_1}(X_i-\overline{X})^2+\sum_{i=1}^{n_2}(Y_i-\overline{Y})^2\right]$是 σ^2 的无偏估计.

解 （1）因为$E(\overline{X})=E\left(\frac{1}{n_1}\sum_{i=1}^{n_1}X_i\right)=\frac{1}{n_1}n_1\mu_1=\mu_1$，

$$E(\overline{Y})=E\left(\frac{1}{n_2}\sum_{i=1}^{n_2}Y_i\right)=\frac{1}{n_2}n_2\mu_2=\mu_2,$$

所以 $$E(\overline{X}-\overline{Y})=E(\overline{X})-E(\overline{Y})=\mu_1-\mu_2,$$

即 $\overline{X}-\overline{Y}$ 是参数 $\mu_1-\mu_2$ 的无偏估计量.

（2）因为 $$\sum_{i=1}^{n_1}(X_i-\overline{X})^2=(n_1-1)S_1^2,$$

$$\sum_{i=1}^{n_2}(Y_i-\overline{Y})^2=(n_2-1)S_2^2,$$

其中，S_1^2 为 $X_1,X_2,\cdots,X_{n_1}$的样本方差，S_2^2 为 $Y_1,Y_2,\cdots,Y_{n_2}$的样本方差，且

$$E\left[\sum_{i=1}^{n_1}(X_i-\overline{X})^2\right]=(n_1-1)E(S_1^2)=(n_1-1)\sigma^2,$$

$$E\left[\sum_{i=1}^{n_2}(Y_i-\overline{Y})^2\right]=(n_2-1)E(S_2^2)=(n_2-1)\sigma^2,$$

所以 $$E\left[\sum_{i=1}^{n_1}(X_i-\overline{X})^2+\sum_{i=1}^{n_2}(Y_i-\overline{Y})^2\right]=(n_1+n_2-2)\sigma^2,$$

故 $E(S_W^2)=\sigma^2$，即 S_W^2 是 σ^2 的无偏估计量.

例28 设 $X_1,X_2,\cdots,X_n$ 是 $X\sim N(0,\sigma^2)$ $(\sigma^2>0)$的一个样本，证明：$\hat{\sigma}=\frac{1}{n}\sqrt{\frac{\pi}{2}}\sum_{i=1}^{n}|X_i|$是 σ 的无偏估计.

证 先计算 $E(|X_i|)$，有

$$E(|X_i|)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}\frac{|x|}{\sigma}\mathrm{e}^{-x^2/(2\sigma^2)}\mathrm{d}x=\sqrt{\frac{2}{\pi}}\int_{0}^{+\infty}\frac{x}{\sigma}\mathrm{e}^{-x^2/(2\sigma^2)}\mathrm{d}x$$

$$=\sqrt{\frac{2}{\pi}}\sigma\left[-e^{-x^2/(2\sigma^2)}\right]\Bigg|_0^{+\infty}=\sqrt{\frac{2}{\pi}}\sigma.$$

所以 $E(\hat{\sigma})=\frac{1}{n}\sqrt{\frac{\pi}{2}}\sum_{i=1}^{n}E(|X_i|)=\frac{1}{n}\sqrt{\frac{\pi}{2}}\cdot\sqrt{\frac{2}{\pi}}\cdot n\sigma=\sigma.$

即 $\hat{\sigma}$ 是 σ 的无偏估计量.

例 29 设 $X_1,X_2,\cdots,X_n$ 是总体 $U(0,\theta)$ 的一个样本,证明:$\hat{\theta}_1=2\overline{X}$ 和 $\hat{\theta}_2=\frac{n+1}{n}X_{(n)}$ 是 θ 的一致估计,$X_{(n)}=\max\{X_i\}$.

证 由例 22 知

$$E(\hat{\theta}_1)=\theta,\quad D(\hat{\theta}_1)=\frac{\theta^2}{3n},$$

$$E(\hat{\theta}_2)=\theta,\quad D(\hat{\theta}_2)=\frac{\theta}{n(n+2)},$$

依契比雪夫不等式,对任给的 $\varepsilon>0$,当 $n\to\infty$ 时,有

$$P\{|\hat{\theta}_1-\theta|\geqslant\varepsilon\}\leqslant\frac{D(\hat{\theta}_1)}{\varepsilon^2}=\frac{\theta^2}{3n\varepsilon^2}\longrightarrow 0,$$

$$P\{|\hat{\theta}_2-\theta|\geqslant\varepsilon\}\leqslant\frac{D(\hat{\theta}_2)}{\varepsilon^2}=\frac{\theta^2}{n(n+2)\varepsilon^2}\longrightarrow 0,$$

所以,$\hat{\theta}_1$ 和 $\hat{\theta}_2$ 都是 θ 的一致估计量.

例30 设估计量 $\hat{\theta}=\hat{\theta}_n=\hat{\theta}(X_1,X_2,\cdots,X_n)$ 是 θ 的估计量,满足 $\lim\limits_{n\to\infty}E[(\hat{\theta}_n-\theta)^2]=0$,证明:$\hat{\theta}_n$ 是 θ 的一致估计量.

证 首先证明 $P\{|\hat{\theta}_n-\theta|\geqslant\varepsilon\}\leqslant\frac{E[(\hat{\theta}_n-\theta)^2]}{\varepsilon^2}$. 设 $\hat{\theta}=\hat{\theta}_n$ 的概率密度为 $f_n(x)$,于是

$$P\{|\hat{\theta}_n-\theta|\geqslant\varepsilon\}=\int_{|x-\theta|\geqslant\varepsilon}f_n(x)\mathrm{d}x\leqslant\int_{|x-\theta|\geqslant\varepsilon}\frac{(x-\theta)^2}{\varepsilon^2}f_n(x)\mathrm{d}x$$
$$\leqslant\frac{1}{\varepsilon^2}\int_{-\infty}^{+\infty}(x-\theta)^2f_n(x)\mathrm{d}x=\frac{1}{\varepsilon^2}E[(\hat{\theta}_n-\theta)^2].$$

由 $\lim\limits_{n\to\infty}E[(\hat{\theta}_n-\theta)^2]=0$ 知,$\lim\limits_{n\to\infty}P\{|\hat{\theta}_n-\theta|\geqslant\varepsilon\}=0$,即 $\hat{\theta}_n$ 是 θ 的一致估计量.

例 31 设总体$X\sim U(1,\theta)$，求θ的矩估计量$\hat{\theta}$，并证明$\hat{\theta}$是θ的一致估计量.

证 先求出θ的矩估计. 因为X的概率密度函数为

$$f(x)=\begin{cases}1/(\theta-1), & 1<x<\theta,\\ 0, & \text{其它},\end{cases}$$

所以$E(X)=\dfrac{1+\theta}{2}$. 令$\overline{X}=E(X)$，得θ的矩估计量$\hat{\theta}=2\overline{X}-1$.

又
$$E(X)=\frac{1+\theta}{2},\quad E(\overline{X})=\frac{1+\theta}{2},$$
有
$$E(\hat{\theta})=E(2\overline{X}-1)=2E(\overline{X})-1=1+\theta-1=\theta,$$
所以，$\hat{\theta}$是θ的无偏估计量.

又
$$D(\overline{X})=\frac{1}{n}D(X)=\frac{(\theta-1)^2}{12n}\longrightarrow 0,\quad n\to\infty,$$
所以$\overline{X}$是$\dfrac{1+\theta}{2}$的一致估计量.

令$g(X)=2X-1$，则$g\left(\dfrac{1+\theta}{2}\right)=\theta$. 因为$g(X)$在$\dfrac{1+\theta}{2}$连续，故$g(\overline{X})=2\overline{X}-1$是$g\left(\dfrac{1+\theta}{2}\right)=\theta$的一致估计.

例 32 设对总体X，$E(X)$，$D(X)$存在，$X_1,X_2,\cdots,X_n$是X的一个样本，证明：样本均值$\overline{X}=\dfrac{1}{n}\sum\limits_{i=1}^{n}X_i$为总体均值$E(X)$的一致无偏估计量.

证 因为
$$E(\overline{X})=E\left(\frac{1}{n}\sum_{i=1}^{n}X_i\right)=\frac{1}{n}\sum_{i=1}^{n}E(X_i)=E(X),$$
所以，$\overline{X}$是$E(X)$的无偏估计量.

又由契比雪夫大数定律知，对任意$\varepsilon>0$，当$n\to\infty$时，有
$$\lim_{n\to\infty}P\left\{\left|\frac{1}{n}\sum_{i=1}^{n}X_i-E(X)\right|\geqslant\varepsilon\right\}=0,$$
故$\overline{X}$是$E(X)$的一致估计量，从而$\overline{X}$是$E(X)$的一致无偏估计量.

例 33 设抽得X_1是$X\sim U(0,\theta)$的一个样本，试证明$\hat{\theta}_1=2X_1$

和 $\hat{\theta}_2=X_1$ 都不是 θ 的一致估计量.

证 因 $E(\hat{\theta}_1)=2E(X_1)=2\times\dfrac{\theta-0}{2}=\theta$,故 $\hat{\theta}_1$ 是 θ 的无偏估计.

又因 $E(\hat{\theta}_2)=E(X_1)=\dfrac{\theta-0}{2}=\dfrac{\theta}{2}\neq\theta$,故 $\hat{\theta}_2$ 是 θ 的有偏估计.

对于任意的 $\varepsilon>0$,因为 $|\hat{\theta}_1-\theta|=|2X_1-\theta|$,所以 $|2X-\theta|\geqslant\varepsilon$ 等价于 $2X_1-\theta\geqslant\varepsilon$ 和 $2X_1-\theta\leqslant-\varepsilon$,即

$$X_1\geqslant(\varepsilon+\theta)/2>0 \quad 和 \quad X_1\leqslant\theta-\varepsilon/2.$$

又因为 $X\sim U(0,\theta)$,所以 $f(x)=\dfrac{1}{\theta}$ $(0<x<\theta)$,于是

$$\begin{aligned}P\{|\hat{\theta}_1-\theta|\}\geqslant\varepsilon&=P\{|2X_1-\theta|\geqslant\varepsilon\}=\int_{|2x_1-\theta|\geqslant\varepsilon}f(x)\mathrm{d}x\\&=\int_{(\theta+\varepsilon)/2}^{0}\frac{1}{\theta}\mathrm{d}x+\int_{0}^{(\theta-\varepsilon)/2}\frac{1}{\theta}\mathrm{d}x\\&=\frac{\theta-\varepsilon}{0}\not\to 0\ (n\to\infty).\end{aligned}$$

所以,$\hat{\theta}_1$ 不是 θ 的一致估计量.

类似可证 $\hat{\theta}_2$ 也不是 θ 的一致估计量.

通过上述例题可以看出,一致性的证明方法比较多,可以依据定义直接讨论 $\lim\limits_{n\to\infty}\hat{\theta}$,或者 $\lim\limits_{n\to\infty}P\{|\hat{\theta}-\theta|\geqslant\varepsilon\}$,而解题的技巧就在对 $P\{|\hat{\theta}-\theta|\geqslant\varepsilon\}$ 的处理上;也可以用契比雪夫不等式证明,这时要求出 $D(\hat{\theta})$;在存在随机变量序列的情形,还可以用大数定律来证. 读者要学会判别条件,选择恰当的方法进行证明.

第二节 区间估计

主要内容

用以一定的概率包含 θ 的真值的区间来估计未知参数的方法

称为区间估计，所得到的区间称为置信区间.

设总体X的分布$F(x;\theta)$中含有未知参数θ.对于给定的α值$(0<\alpha<1)$，若存在X的两个统计量$\underline{\theta}=\underline{\theta}(X_1,X_2,\cdots,X_n)$和$\overline{\theta}=\overline{\theta}(X_1,X_2,\cdots,X_n)$，使得

$$P\{\underline{\theta}(X_1,X_2,\cdots,X_n)<\theta<\overline{\theta}(X_1,X_2,\cdots,X_n)\}=1-\alpha,$$

则称随机区间$(\underline{\theta},\overline{\theta})$为$\theta$的置信度为$1-\alpha$的双侧置信区间. $\underline{\theta}$和$\overline{\theta}$分别称为θ的置信度为$1-\alpha$的双侧置信区间的置信下限和置信上限，$1-\alpha$为置信度. 若只考虑

$$P\{\underline{\theta}(X_1,X_2,\cdots,X_n)<\theta\}=1-\alpha$$

或
$$P\{\theta<\overline{\theta}(X_1,X_2,\cdots,X_n)\}=1-\alpha,$$

则称随机区间$(\underline{\theta},+\infty)$或$(-\infty,\overline{\theta})$为$\theta$的置信度为$1-\alpha$的单侧置信区间. $\underline{\theta}$和$\overline{\theta}$分别称为单侧置信下限和单侧置信上限.

一、单个正态总体均值与方差的区间估计

$X\sim N(\mu,\sigma^2)$，$X_1,X_2,\cdots,X_n$为X的一个样本，$\overline{X}$和S^2分别为样本均值与样本方差.

1. 均值μ的置信区间　$X\sim N(\mu,\sigma^2)$

(1) σ^2已知时，$\overline{X}\sim N\left(\mu,\dfrac{\sigma^2}{n}\right)$. 选用估计量$U=\dfrac{\overline{X}-\mu}{\sigma/\sqrt{n}}\sim N(0,1)$，则$\mu$的置信度为$1-\alpha$的置信区间为

$$\left(\overline{X}-\frac{\sigma}{\sqrt{n}}Z_{\alpha/2},\overline{X}+\frac{\sigma}{\sqrt{n}}Z_{\alpha/2}\right),$$

$Z_{\alpha/2}$是标准正态分布的上$\alpha/2$分位点.

(2) σ^2未知时，选用估计量$T=\dfrac{\overline{X}-\mu}{S/\sqrt{n}}\sim t(n-1)$，则$\mu$的置信度为$1-\alpha$的置信区间为

$$\left(\overline{X}-\frac{S}{\sqrt{n}}t_{\alpha/2}(n-1),\overline{X}+\frac{S}{\sqrt{n}}t_{\alpha/2}(n-1)\right).$$

(3) 估计μ时，样本容量n的确定.

若σ^2已知，要求μ的置信度为$1-\alpha$的置信区间长度不超过l，

则 $n \geqslant (2\sigma Z_{\alpha/2}/l)^2$.

若 σ^2 未知，而 n 较大时，可由经验估计出 σ_0^2，则 $n \geqslant (2\sigma_0 Z_{\alpha/2}/l)^2$.

2. 方差 σ^2 的置信区间

若 μ 未知，选用估计量 $\chi^2 = \frac{(n-1)}{\sigma^2}S^2 \sim \chi^2(n-1)$，则 σ^2 的置信度为 $1-\alpha$ 的置信区间为

$$\left(\frac{(n-1)S^2}{\chi^2_{\alpha/2}(n-1)}, \frac{(n-1)S^2}{\chi^2_{1-\alpha/2}(n-1)}\right).$$

若 μ 已知，利用 σ^2 的极大似然估计是 $\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i-\mu)^2$，得 σ^2 的置信度为 $1-\alpha$ 的置信区间为

$$\left(\frac{\sum_{i=1}^{n}(X_i-\mu)^2}{\chi^2_{\alpha/2}(n)}, \frac{\sum_{i=1}^{n}(X_i-\mu)^2}{\chi^2_{1-\alpha/2}(n)}\right).$$

二、两个正态总体均值差与方差比的区间估计

1. 两个正态总体均值差的区间估计

总体 $X \sim N(\mu_1, \sigma_1^2)$，总体 $Y \sim N(\mu_2, \sigma_2^2)$，$X_1, X_2, \cdots, X_{n_1}$ 是 X 的一个样本，$Y_1, Y_2, \cdots, Y_{n_2}$ 是 Y 的一个样本. $\overline{X}, \overline{Y}$ 和 S_1^2, S_2^2 分别是 X，Y 的样本均值和样本方差.

(1) σ_1^2, σ_2^2 均已知，则 $\overline{X}-\overline{Y}$ 是 $\mu_1-\mu_2$ 的无偏估计，$\overline{X}-\overline{Y} \sim N\left(\mu_1-\mu_2, \frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}\right)$. 选用估计量

$$U = \frac{(\overline{X}-\overline{Y})-(\mu_1-\mu_2)}{\sqrt{\sigma_1^2/n_1+\sigma_2^2/n_2}} \sim N(0,1),$$

则 $\mu_1-\mu_2$ 的置信度为 $1-\alpha$ 的置信区间为

$$\left(\overline{X}-\overline{Y}-Z_{\alpha/2}\sqrt{\sigma_1^2/n_1+\sigma_2^2/n_2}, \overline{X}-\overline{Y}+Z_{\alpha/2}\sqrt{\sigma_1^2/n_1+\sigma_2^2/n_2}\right).$$

(2) $\sigma_1^2=\sigma_2^2=\sigma^2$，但 σ^2 未知，选用估计量

$$T = \frac{(\overline{X}-\overline{Y})-(\mu_1-\mu_2)}{S_W\sqrt{1/n_1+1/n_2}} \sim t(n_1+n_2-2),$$

其中

$$S_W=\frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2},$$

则 $\mu_1-\mu_2$ 的置信度为 $1-\alpha$ 的置信区间为

$$\left(\overline{X}-\overline{Y}\mp t_{\alpha/2}(n_1+n_2-2)S_W\sqrt{1/n_1+1/n_2}\right).$$

(3) σ_1^2,σ_2^2 未知，但为大样本情形，则 $\mu_1-\mu_2$ 的置信度为 $1-\alpha$ 的近似置信区间为

$$\left(\overline{X}-\overline{Y}-Z_{\alpha/2}\sqrt{S_1^2/n_1+S_2^2/n_2},\overline{X}-\overline{Y}+Z_{\alpha/2}\sqrt{S_1^2/n_1+S_2^2/n_2}\right).$$

2. 两个正态总体方差比的置信区间

$X\sim N(\mu_1,\sigma_1^2)$，$Y\sim N(\mu_2,\sigma_2^2)$，$X_1,X_2,\cdots,X_{n_1}$ 为 X 的一个样本，$Y_1,Y_2,\cdots,Y_{n_2}$ 为 Y 的一个样本，S_1^2,S_2^2 分别为两样本的样本方差，$\mu_1,\mu_2,\sigma_1^2,\sigma_2^2$ 均未知. 选用估计量

$$F=\frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2}\sim F(n_1-1,n_2-1),$$

则 σ_1^2/σ_2^2 的置信度为 $1-\alpha$ 的置信区间为

$$\left(\frac{S_1^2}{S_2^2F_{\alpha/2}(n_1-1,n_2-1)},\frac{S_1^2}{S_2^2F_{1-\alpha/2}(n_1-1,n_2-1)}\right).$$

疑 难 解 析

1. 什么是区间估计？有了点估计为什么还要引入区间估计？

答 用以一定概率 $(1-\alpha)$ 包含真值 θ 的区间来估计未知参数的方法称为区间估计. 点估计是利用样本值求得参数 θ 的一个近似值来估计未知参数 θ，但不知近似的精确程度和可信程度，因此虽有一定参考价值，但毕竟实用意义不大. 区间估计则通过两个统计量 $\underline{\theta}$ 和 $\overline{\theta}$，确定了一个随机区间 $(\underline{\theta},\overline{\theta})$，使得该区间内包含真值 θ 的概率不小于 $1-\alpha$. 区间估计不仅提供了 θ 的一个估计范围，还给出了估计的精度与可信程度，有广泛的实用意义.

2. 置信度 $1-\alpha$ 的意义是什么?

答 置信度有两种理解方式.

对于一个置信区间 $(\underline{\theta},\overline{\theta})$ 而言, $1-\alpha$ $(0<\alpha<1)$ 表示可信程度,即随机区间 $(\underline{\theta},\overline{\theta})$ 中包含未知参数 θ 的概率不小于事先设定的 $1-\alpha$.

对于区间估计的设计而言, $1-\alpha$ 又表示在样本容量不变的情况下,在反复抽样所得到的全部区间中,包含 θ 真值的区间不小于 $100(1-\alpha)\%$.

一般地, $1-\alpha$ 值越大,由样本值所得的区间 $(\underline{\theta},\overline{\theta})$ 覆盖 θ 的置信度越大. 而 $(\underline{\theta},\overline{\theta})$ 的长度越小,又反映估计 θ 的精度越高. 在 n 一定的情况下,精度与置信度不可能兼得. 建立区间估计理论的著名统计学家 Neyman 提出的原则是,先照顾可靠程度,即置信度优于精度,在满足 $P\{\underline{\theta}<\theta<\overline{\theta}\}=1-\alpha$ 的前提下,使精度尽可能地高.

3. 进行区间估计的一般步骤有哪些?

答 区间估计的一般步骤如下.

(1) 根据实际问题的条件,确定未知参数的一个估计量 $Z=Z(X_1,X_2,\cdots,X_n;\theta)$,不含 θ 外的其它未知参数,且 Z 的分布已知.

(2) 对于事先给定的置信度 $1-\alpha$,确定常数 a,b,使 $P\{a<Z(X_1,X_2,\cdots,X_n;\theta)<b\}=1-\alpha$.

(3) 求出与 $a<Z(X_1,X_2,\cdots,X_n;\theta)<b$ 等价的不等式 $\underline{\theta}<\theta<\overline{\theta}$,则 $(\underline{\theta},\overline{\theta})$ 即为所求的 θ 的置信度为 $1-\alpha$ 的置信区间.

事实上,使 $P\{a<Z(X_1,X_2,\cdots,X_n;\theta)<b\}=1-\alpha$ 的数 a,b 有无穷多组,所以置信区间是不唯一的. 但一般总是选择对称形式或近似对称形式的置信区间,这是为了计算的方便.

4. 怎样评价两个正态总体下区间估计的结果?

答 对于两个总体均值差的区间估计,若 $(\underline{\theta},\overline{\theta})$ 包含数零,可以认为两个总体的均值没有大的差异;若置信下限大于零,可以认为 μ_1 大于 μ_2;若置信上限小于零,可以认为 μ_1 小于 μ_2.

对两个正态总体方差比的区间估计，若$(\underline{\theta},\overline{\theta})$包含1，可以认为两个总体的方差没有大的区别；若置信下限大于1，可以认为σ_1^2大于σ_2^2；若置信上限小于1，可以认为σ_1^2小于σ_2^2.

方法、技巧与典型例题分析

对于实际问题求参数的区间估计，首先是要认真分析问题的条件，确定恰当的估计量，选择好区间估计的形式，然后直接计算即可；其次是分清问题是单侧置信区间还是双侧置信区间，以便确定用α还是$\alpha/2$.

一、单个正态总体均值与方差的区间估计

例1 已知某地幼儿的身高服从正态分布. 现从该地一幼儿园的大班抽查了9名幼儿，测得身高(单位:cm)分别为115，120，131，115，109，115，115，105，110. 设大班幼儿身高总体的标准差$\sigma=7$ cm，在$\alpha=0.05$下，求总体均值μ的置信区间.

解 已知$\sigma^2=7^2$，选用估计量$U=\dfrac{\overline{X}-\mu}{\sigma/\sqrt{n}}$，又$n=9$，$\bar{x}=115$，$Z_{0.025}=1.96$，所以$\mu$的置信度为0.95的置信区间为

$$\left(115-1.96\times\frac{7}{\sqrt{9}},115+1.96\times\frac{7}{\sqrt{9}}\right)=(110.43,119.57).$$

例2 设X方差为1，样本容量$n=100$，$\bar{x}=5$，求μ的置信度为0.95的置信区间.

解 $n=100$，是大样本. 由中心极限定理有，$\overline{X}\dot{\sim}N(\mu,\sigma^2/n)$，所以选用估计量$U$. μ的置信度为0.95的置信区间为

$$\left(5-\frac{1}{10}\times1.96,5+\frac{1}{10}\times1.96\right)=(4.804,5.196).$$

例3 从一大批电子管中随机抽取了100只，抽取的电子管的平均寿命为1000 h. 设电子管寿命服从正态分布，均方差$\sigma=40$，以置信度0.95求出这批电子管平均寿命μ的置信区间.

解 σ^2 已知，选用估计量 U. 又 $n=100,\sigma=40,\bar{x}=1000,Z_{0.025}=1.96$，所以 μ 的置信度为 0.95 的置信区间为

$$\left(1000-\frac{40}{\sqrt{100}}\times 1.96,1000+\frac{40}{\sqrt{100}}\times 1.96\right)$$
$$=(992.16,1007.84).$$

例 4 为了估计产品使用寿命的均值 μ 和标准差 σ，测试了 10 件产品，求得 $\bar{x}=1500,S=20$. 若已知产品使用寿命服从正态分布 $N(\mu,\sigma^2)$，求出 μ 和 σ^2 的置信度为 0.95 的置信区间.

解 (1) σ^2 未知，选用估计量 $T=\dfrac{\overline{X}-\mu}{S/\sqrt{n}}\sim t(n-1)$. 又 $\bar{x}=1500,S=20,n=10,t_{0.025}(9)=2.2622$，所以 μ 的置信度为 0.95 的置信区间为

$$\left(1500-\frac{20}{\sqrt{10}}\times 2.2622,1500+\frac{20}{\sqrt{10}}\times 2.2622\right)$$
$$=(1485.7,1514.3).$$

(2) μ 未知，选用估计量 $\chi^2=\dfrac{(n-1)S^2}{\sigma^2}\sim\chi^2(n-1)$. 又 $n=10,S=20,\chi^2_{0.025}(9)=19,\chi^2_{0.975}(9)=2.7$，所以 σ^2 的置信度为 0.95 的置信区间为

$$\left(\frac{9\times 400}{19},\frac{9\times 400}{2.7}\right)=(189.47,1333.33).$$

例5 设炮弹的速度 v 服从正态分布，抽取9 发炮弹试验，测得样本方差 $S^2=11$，求炮弹速度 v 的方差 σ^2 与标准差 σ 的置信度为 0.90的置信区间.

解 μ 未知，选用估计量 $\chi^2=\dfrac{(n-1)S^2}{\sigma^2}\sim\chi^2(n-1)$. 又 $n=9,S^2=11,\chi^2_{0.05}(8)=15.507,\chi^2_{0.95}(8)=2.733$，所以 σ^2 的置信度为 0.90 的置信区间为

$$\left(\frac{(9-1)\times 11}{15.507},\frac{(9-1)\times 11}{2.733}\right)=(5.675,32.199);$$

σ 的置信度为 0.90 的置信区间为

$$\left(\sqrt{\frac{(9-1)\times 11}{15.507}},\sqrt{\frac{(9-1)\times 11}{2.733}}\right)=(2.382,5.674).$$

例6 设制造某种产品每件所需时间服从正态分布，现随机记录了5件产品所用工时：10.5,11,11.2,12.5,12.8. 求 μ 的置信度为0.95 的单侧置信上限.

解 σ^2 未知，选用估计量 $T=\frac{\overline{X}-\mu}{S/\sqrt{n}}\sim t(n-1)$. 又 $n=5,\bar{x}=11.6,S^2=0.995,t_{0.05}(5-1)=2.1318$，所以 μ 的置信度为0.95 的单侧置信区间为

$$\left(0,11.6+\frac{0.9975}{\sqrt{5}}\times 2.1318\right)=(0,12.55).$$

故 μ 的置信度为0.95 的单侧置信上限为12.55.

注意，此时的置信下限不一定取 $-\infty$，可视问题确定. 如本题，因工时无负值，故取置信下限为零.

例7 从自动机床加工的9个零件中测得零件的平均长度（单位：mm）为21.4. 设零件长度服从正态分布，求零件长度的均值 μ 的置信度为0.95 的置信区间. 如果：(1) $\sigma=0.15$；(2) σ 未知.

解 (1) 已知 $\sigma=0.15$，选用估计量 $U=\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}$. 又 $\bar{x}=21.4$，$n=9,Z_{0.025}=1.96$，所以 μ 的置信度为0.95 的置信区间为

$$\left(21.4-1.96\times\frac{0.15}{\sqrt{9}},21.4+1.96\times\frac{0.15}{\sqrt{9}}\right)$$
$$=(21.302,21.498).$$

(2) σ 未知，选用估计量 $T=\frac{\overline{X}-\mu}{S/\sqrt{n}}\sim t(n-1)$. 又 $\bar{x}=21.4,n=9,S=S,t_{0.025}(8)=2.306$，所以 μ 的置信度为0.95 的置信区间为

$$\left(21.4-2.306\times\frac{S}{3},21.4+2.306\times\frac{S}{3}\right).$$

例8 求上题中零件长度（单位：mm）的方差 σ^2 的置信度为

0.95的置信区间.如果:(1) $\mu=21.42$;(2) μ 未知.

解 (1) μ 已知时,σ^2 的置信度为 $1-\alpha$ 的置信区间为

$$\left(\frac{\sum_{i=1}^{n}(X_i-\mu)^2}{\chi^2_{\alpha/2}(n)},\frac{\sum_{i=1}^{n}(X_i-\mu)^2}{\chi^2_{1-\alpha/2}(n)}\right).$$

可以算得,$\sum_{i=1}^{n}(X_i-\mu)^2=0.2636$,而 $\chi^2_{0.025}(9)=19.023$,$\chi^2_{0.975}(9)=2.7$,则 σ^2 的置信度为 0.95 的置信区间为

$$\left(\frac{0.2636}{19.023},\frac{0.2636}{2.7}\right)=(0.0139,0.0976).$$

(2) μ 未知时,σ^2 的置信度为 $1-\alpha$ 的置信区间为

$$\left(\frac{(n-1)S^2}{\chi^2_{\alpha/2}(n-1)},\frac{(n-1)S^2}{\chi^2_{1-\alpha/2}(n-1)}\right).$$

又 $\chi^2_{0.025}(8)=17.535$,$\chi^2_{0.975}(8)=2.18$,所以 σ^2 的置信度为0.95的置信区间为

$$\left(\frac{8S^2}{17.535},\frac{8S^2}{2.18}\right).$$

例9 设 S 是总体 $X\sim N(\mu,\sigma^2)$ 的随机样本 $X_1,X_2,\cdots,X_n$ 的方差,μ,σ^2 均未知,问:a,b $(0<a<b)$ 为何值时,σ^2 的 0.95 的置信区间 $\left(\frac{(n-1)S^2}{b}<\sigma^2<\frac{(n-1)S^2}{a}\right)$ 的长度最短?

解 因为
$$\frac{(n-1)S^2}{\sigma^2}\sim\chi^2(n-1),$$
所以
$$P\left\{\frac{(n-1)S^2}{b}<\sigma^2<\frac{(n-1)S^2}{a}\right\}=0.95,$$
而 μ 未知时,σ^2 的置信区间的长度为

$$l=\left(\frac{1}{a}-\frac{1}{b}\right)(n-1)S^2,$$

又
$$P\left\{\frac{(n-1)S^2}{b}<\sigma^2<\frac{(n-1)S^2}{a}\right\}=P\left\{a<\frac{(n-1)S^2}{\sigma^2}<b\right\}$$
$$=\int_a^b f(y)\mathrm{d}y=F(b)-F(a)$$

($f(y)$是$\chi^2(n-1)$的概率密度函数).

要使l达到最小,利用求极值方法得

$$l'_a=\left(-\frac{1}{a^2}+\frac{1}{b^2}b'\right)(n-1)S^2,$$

令上式等于零,解得　　$b^2=a^2b'$.

再对$F(b)-F(a)=0$求关于a的导数,得

$$F'(b)b'-F'(a)=0,\quad 即\quad f(b)b'-f(a)=0\Longrightarrow b'=\frac{f(a)}{f(b)},$$

所以 $b^2=a^2\dfrac{f(a)}{f(b)}$,即当$a,b$满足$b^2f(b)=a^2f(a)$时,区间$\left(\dfrac{(n-1)S^2}{b},\dfrac{(n-1)S^2}{a}\right)$最短.

例10　设总体$X\sim N(\mu,\sigma^2)$,已知$\sigma=\sigma_0$,要使μ的置信度为$1-\alpha$的置信区间长度不大于l,问:应抽取多大容量的样本?

解　因为$\dfrac{\overline{X}-\mu}{\sigma_0/\sqrt{n}}\sim N(0,1)$,所以$\mu$的置信度为$1-\alpha$的置信区间为$(\overline{X}\mp\sigma_0/\sqrt{n}\cdot Z_{\alpha/2})$,所以区间长度为

$$l=2Z_{\alpha/2}\sigma_0/\sqrt{n}.$$

要使$l\leqslant 2Z_{\alpha/2}\sigma_0/\sqrt{n}$,则应有

$$\sqrt{n}\geqslant 2\sigma_0Z_{\alpha/2}/l,\quad 即\quad n\geqslant(2\sigma_0Z_{\alpha/2}/l)^2.$$

例11　设$X_1,X_2,\cdots,X_n$是总体$X\sim N(\mu,\sigma^2)$的一个样本,μ,σ^2均未知,求关于μ的置信度为$1-\alpha$的置信区间的长度l平方的数学期望.

解　在σ^2未知时,选用估计量$T=\dfrac{\overline{X}-\mu}{S/\sqrt{n}}\sim t(n-1)$,$\mu$的置信度为$1-\alpha$的置信区间为

$$(\overline{X}\mp t_{\alpha/2}(n-1)\cdot S/\sqrt{n}),$$

置信区间的长度为

$$l=2t_{\alpha/2}(n-1)\cdot S/\sqrt{n}.$$

$$E(l^2)=E[4t_{\alpha/2}^2(n-1)\cdot S/n]=\frac{4}{n}t_{\alpha/2}^2(n-1)\cdot E(S^2)$$

$$=\frac{4}{n}\sigma^2 t_{\alpha/2}^2(n-1).$$

例12 设$X_1,X_2,\cdots,X_n$是总体$X\sim N(\mu,1)$的一个样本，μ未知，要得到μ的一个长度不超过0.2、置信度为0.99的置信区间，样本容量至少应为多大？

解 因为$\frac{\overline{X}-\mu}{1/\sqrt{n}}\sim N(0,1)$，所以$\mu$的置信度为$1-\alpha$的置信区间为$(\overline{X}\mp Z_{\alpha/2}/\sqrt{n})$.

由$1-\alpha=0.99$，$Z_{\alpha/2}=Z_{0.005}=2.576$知，置信区间的长度为

$$l=2Z_{\alpha/2}/\sqrt{n}=2\times 2.576/\sqrt{n}.$$

要使$l\leqslant 0.2$，则应有$n\geqslant(2\times 2.576/0.2)^2$，即

$$n\geqslant\left(\frac{5.152}{0.2}\right)^2=663.5776.$$

所以，样本容量至少应取664.

例13 在测量反应时间中，一心理学家估计的标准差是0.05(单位：s)．为了以0.95的置信度使他对平均反应时间的估计的误差不超过0.01 s，问：应取多大的样本容量？

解 以X记反应时间，则$\mu=E(X)$表示平均反应时间，$S=0.05$；当n充分大时，有

$$T=\frac{\overline{X}-\mu}{S/\sqrt{n}}\sim N(0,1),$$

故要求样本容量n满足

$$P\{|\overline{X}-\mu|\leqslant 0.01\}=P\left\{\frac{|\overline{X}-\mu|}{S/\sqrt{n}}\leqslant\frac{0.01\sqrt{n}}{0.05}\right\}=0.95.$$

由$Z_{0.025}=1.96$，得

$$\sqrt{n}\approx\frac{0.05}{0.01}\times 1.96=9.8\Longrightarrow n\geqslant 96.04,$$

所以，样本容量$n\geqslant 96$.

二、两个总体均值差与方差比的区间估计

两个总体问题在生产、科学技术及管理上都有较广泛的应用.解题的方法是:认真分析具体问题的条件,选择适当的估计量,熟练运用公式,正确进行计算.注意运算的技巧,以使解题过程更为简捷.

例14 设超大牵伸纺机所纺纱的抗拉强度$X\sim N(\mu_1,2.18^2)$,普通纺机所纺纱的抗拉强度$Y\sim N(\mu_2,1.76^2)$.现对前者抽取一容量$n_1=200$的样本,对后者抽取一容量$n_2=100$的样本.经计算,得$\bar{x}=5.32$,$\bar{y}=5.76$.求$\mu_1-\mu_2$的置信度为0.95的置信区间.

解 σ_1^2,σ_2^2为已知,选用估计量U.因为$n_1=200,n_2=100,\sigma_1^2=2.18^2,\sigma_2^2=1.76^2,\bar{x}=5.32,\bar{y}=5.76,\dfrac{\sigma_1^2}{n_1}+\dfrac{\sigma_2^2}{n_2}=0.0547,\sqrt{\dfrac{\sigma_1^2}{n_1}+\dfrac{\sigma_2^2}{n_2}}=0.234$,所以,$\mu_1-\mu_2$的置信度为0.95的置信区间为

$$\left(\overline{X}-\overline{Y}-Z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}},\overline{X}-\overline{Y}+Z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}\right)$$
$$=(-0.44-1.96\times0.234,-0.44+1.96\times0.234)$$
$$=(-0.899,0.0186).$$

例15 随机地从A批导线中抽取4根,又从B批导数中抽取5根,测得电阻数据(单位:Ω)为

A批:0.148, 0.142, 0.143, 0.137,

B批:0.140, 0.142, 0.136, 0.138, 0.140.

设A批电阻$X\sim N(\mu_1,\sigma^2)$,B批电阻$Y\sim N(\mu_2,\sigma^2)$,两个样本相互独立.又μ_1,μ_2,σ^2均未知,求$\mu_1-\mu_2$的置信度为0.95的置信区间.

解 $\sigma_1^2=\sigma_2^2=\sigma^2$,但$\sigma^2$未知,选用估计量$T$.经计算,$\bar{x}=0.1413,\bar{y}=0.1392,S_1^2=8.25\times10^{-6},S_2^2=5.2\times10^{-5},S_W=2.55\times10^{-3},t_{0.025}(4+5-2)=2.3646$.所以,$\mu$的置信度为0.95的置信区间为

$$\left(\overline{X}-\overline{Y}\mp t_{\alpha/2}(n_1+n_2-2)S_W\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}\right)$$

$=(0.0021\mp 2.3646\times 2.55\times 10^{-3}\times 0.67)$

$=(-0.002,0.006)$.

由于 $\mu_1-\mu_2$ 的置信区间包含零,可以认为两个总体的均值无明显差异.

例16 为了估计磷肥对农作物增产的作用,选取20块条件基本相同的土地,其中10块地施磷肥,另外10块不施磷肥.测得亩(1亩=10000/15 m²)产量(单位:kg)如表7.1所示.设两种情形的亩产量均服从正态分布,且方差相同,试对两种情形下平均亩产量之差作出区间估计($\alpha=0.05$).

表7.1

施磷肥	620	570	650	600	630	580	570	600	600	580
不施肥	560	590	560	570	580	570	600	550	570	550

解 以X记施磷肥的亩产总体,以Y记不施磷肥的亩产总体.$\sigma_1^2=\sigma_2^2=\sigma^2$,但$\sigma^2$未知,选用估计量$T$.

$$n_1=10,\quad n_2=10,\quad t_{0.025}(10+10-2)=t_{0.025}(18)=2.1009,$$

$$\bar{x}=600,\quad \sum_{i=1}^{n}(x_i-\bar{x})^2=(n_1-1)S_1^2=6400,$$

$$\bar{y}=570,\quad \sum_{i=1}^{n}(y_i-\bar{y})^2=(n_2-1)S_2^2=2400,$$

$$S_W=\sqrt{[(n_1-1)S_1^2+(n_2-1)S_2^2]/(n_1+n_2-2)}=22.111,$$

$$\sqrt{1/n_1+1/n_2}=0.447,$$

所以 $\mu_1-\mu_2$ 的置信度为0.95的置信区间为

$$\left(\overline{X}-\overline{Y}\mp t_{\alpha/2}(n_1+n_2-2)S_W\sqrt{1/n_1+1/n_2}\right)$$

$$=(30\mp 2.1009\times 22.111\times 0.447)=(9.236,50.764).$$

可见,施磷肥土地的平均亩产高于不施磷肥的土地的平均亩产量.

例17 生产厂家与使用厂家分别对某种染料的有效含量作了

13 次与10 次测定，测定值的方差分别为$S_1^2=0.7241$，$S_2^2=0.6872$，设两厂的测定值都服从正态分布，其方差分别为σ_1^2和σ_2^2，试求方差比σ_1^2/σ_2^2的置信度为0.90的置信区间.

解 是μ_1和μ_2未知情形的区间估计，选用估计量F.因为$n_1=13$，$n_2=10$，$S_1^2=0.7241$，$S_2^2=0.6872$，$F_{0.05}(13-1,10-1)=F_{0.05}(12,9)=3.09$，$F_{0.95}(12,9)=1/F_{0.05}(9,12)=1/2.8=0.3571$，所以，$\sigma_1^2/\sigma_2^2$的置信度为0.90的置信区间为

$$\left(\frac{S_1^2}{S_2^2F_{0.05}(12,9)},\frac{S_1^2}{S_2^2F_{0.95}(12,9)}\right)=(0.29,2.95).$$

由于置信区间包含1，可以认为两总体的方差没有大的差异.

例18 从甲、乙两厂生产的蓄电池产品中分别抽出一批样品，测得蓄电池的电荷量(单位:A·h,1 A·h=3.6 kC)如表7.2所示.设甲、乙两工厂的蓄电池的电荷量分别服从正态分布$X\sim N(\mu_1,\sigma_1^2)$，$Y\sim N(\mu_2,\sigma_2^2)$.求：

(1) 电荷量的方差比σ_1^2/σ_2^2的置信度为0.95的置信区间；

(2) 电荷量的均值差$(\mu_1-\mu_2)$的置信度为0.95的置信区间(设$\sigma_1^2=\sigma_2^2$).

表7.2

甲厂	144	141	138	142	141	138	143	137		
乙厂	142	143	139	140	138	141	140	138	142	136

解 $\bar{x}=140.5$，$\bar{y}=139.9$，$S_1^2=2.563$，$S_2^2=2.183$.

(1) μ_1,μ_2均未知，选用估计量$F\sim F(n_1-1,n_2-1)$.又$n_1=8$，$n_2=10$，$F_{0.025}(7,9)=4.20$，$F_{0.975}(7,9)=1/F_{0.025}(9,7)=1/4.82=0.2075$，所以$\sigma_1^2/\sigma_2^2$的置信度为0.95的置信区间为

$$\left(\frac{2.563^2}{2.183^2\times4.20},\frac{2.563^2}{2.183^2\times0.2075}\right)=(0.328,6.642).$$

由于置信区间包含1，可以认为两个总体的方差相等，没有明显的

差异.

(2) $\sigma_1^2=\sigma_2^2=\sigma^2$,但 σ^2 未知,选用估计量 $T\sim t(n_1+n_2-2)$. 又

$$n_1=8, n_2=10, \quad \sqrt{1/n_1+1/n_2}=0.474,$$

$$S_W=\sqrt{(7\times 2.563^2+9\times 2.183^2)/16}=2.357,$$

所以 $\mu_1-\mu_2$ 的置信度为 0.95 的置信区间为

$$(\overline{X}-\overline{Y}\mp t_{\alpha/2}(n_1+n_2-2)S_W\sqrt{1/n_1+1/n_2})$$
$$=(0.6\mp 2.1199\times 2.357\times 0.474)=(-1.768, 2.968).$$

由于置信区间包含 0,可认为两个总体的均值没有大的差异.

例 19 设总体 $X\sim N(\mu_1,\sigma_1^2)$,$Y\sim N(\mu_2,\sigma_2^2)$,其中 μ_1,μ_2 已知,σ_1^2,σ_2^2 未知. 从总体 X 和 Y 中分别抽取容量为 n_1 和 n_2 的样本,且两样本相互独立,证明:方差比 σ_1^2/σ_2^2 的置信度为 $1-\alpha$ 的置信区间为

$$\left(\frac{\hat{\sigma}_1^2/\hat{\sigma}_2^2}{F_{\alpha/2}(n_1,n_2)}, \frac{\hat{\sigma}_1^2/\hat{\sigma}_2^2}{F_{1-\alpha/2}(n_1,n_2)}\right),$$

其中 $$\hat{\sigma}_1^2=\frac{1}{n_1}\sum_{i=1}^{n_1}(X_i-\mu_1)^2, \quad \hat{\sigma}_2^2=\frac{1}{n_2}\sum_{i=1}^{n_2}(Y_i-\mu_2)^2.$$

证 因为 $\dfrac{\sum_{i=1}^{n_1}(X_i-\mu_1)^2}{\sigma_1^2}\sim\chi^2(n_1)$, $\dfrac{\sum_{i=1}^{n_2}(Y_i-\mu_2)^2}{\sigma_2^2}\sim\chi^2(n_2)$,

$\chi^2(n_1)$ 与 $\chi^2(n_2)$ 相互独立,所以,依 F 分布定义,有

$$\frac{\hat{\sigma}_1^2}{\sigma_1^2}\bigg/\frac{\hat{\sigma}_2^2}{\sigma_2^2}=\frac{\hat{\sigma}_1^2\sigma_2^2}{\hat{\sigma}_2^2\sigma_1^2}\sim F(n_1,n_2),$$

故 $$P\left\{F_{1-\alpha/2}(n_1,n_2)\leqslant\frac{\hat{\sigma}_1^2\sigma_2^2}{\hat{\sigma}_2^2\sigma_1^2}\leqslant F_{\alpha/2}(n_1,n_2)\right\}=1-\alpha,$$

即 $$P\left\{\frac{\hat{\sigma}_1^2/\hat{\sigma}_2^2}{F_{\alpha/2}(n_2,n_1)}\leqslant\frac{\sigma_1^2}{\sigma_2^2}\leqslant\frac{\hat{\sigma}_1^2/\hat{\sigma}_2^2}{F_{1-\alpha/2}(n_2,n_1)}\right\}=1-\alpha.$$

于是 σ_1^2/σ_2^2 的置信度为 $1-\alpha$ 的置信区间为

$$\left(\frac{\hat{\sigma}_1^2/\hat{\sigma}_2^2}{F_{\alpha/2}(n_2,n_1)}, \frac{\hat{\sigma}_1^2/\hat{\sigma}_2^2}{F_{1-\alpha/2}(n_2,n_1)}\right).$$

第三节 关于总体比例的估计

主要内容

在一些实际问题中，经常需要估计总体中含有某种特征的个体占总体全部个体的比例.这种在一总体中具备某个特征的个体占总体全部个体的比例称为总体比例，用 p 表示.

总体比例问题实际上是一个二项分布问题.

1. p 的点估计

设 $X \sim B(n,p)$，则 p 的点估计 $\hat{p}=\overline{X}/n$.

2. p 的区间估计

根据中心极限定理，当 $n\to\infty$ 时，$\hat{p}\sim N(p,p(1-p)/n)$. 以 $\hat{p}$ 代替 p，得 p 的置信度为 $1-\alpha$ 的(近似)置信区间为

$$(\hat{p}-Z_{\alpha/2}\sqrt{p(1-p)/n},\hat{p}+Z_{\alpha/2}\sqrt{p(1-p)/n}).$$

3. 估计总体比例 p 时，样本容量 n 的确定

当置信区间长度为 l 时，应有

$$n\geqslant\hat{p}(1-\hat{p})(2Z_{\alpha/2}/l)^2.$$

因为 $0\leqslant\hat{p}<1$，$\hat{p}(1-\hat{p})\leqslant 1/4$，所以 $n\geqslant(Z_{\alpha/2}/l)^2$. 当 p 接近于1或0时，取适当的下限或上限 p_0，有

$$n\geqslant p_0(1-p_0)(2Z_{\alpha/2}/l)^2,$$

当 $n<35$ 时，应适当放大.

4. 两个总体比例之差的估计

设有两个总体 X 和 Y，各有 N_1 和 N_2 个个体，且 N_1 和 N_2 都很大，p_1 和 p_2 分别为 X 和 Y 的总体比例. 若从 X 和 Y 中分别抽取 n_1 和 n_2 个个体，而 n_1/N_1 和 n_2/N_2 都很小，其中所含某种性质的个体

比例分别为 $\hat{p}_1$ 和 $\hat{p}_2$.

当n_1,n_2很大,而p_1,p_2不接近于0或1时,$\hat{p}_1-\hat{p}_2$的抽样分布近似服从正态分布,且

$$\mu=p_1-p_2,\quad \sigma^2=p_1(1-p_1)/n_1+p_2(1-p_2)/n_2,$$

p_1-p_2的置信度为$1-\alpha$的置信区间为

$$(\hat{p}_1-\hat{p}_2\mp Z_{\alpha/2}\sqrt{\hat{p}_1(1-\hat{p}_1)/n_1+\hat{p}_2(1-p_2)/n_2}).$$

疑难解析

什么是总体比例问题?

答 在一个总体中,具有某个特征的个体在总体中所占的比例称为总体比例,如:某年级学生中女生的比例,戴近视镜学生的比例;企业中有学位职工的比例;某产品中次品的比例;某小区住户中有太阳能热水器住户的比例;等等.

总体比例问题可以视为二项分布问题.即把具有特征当作“试验成功”,则$X\sim B(n,p)$.当n很大时,依中心极限定理,有

$$(\overline{X}-np)/\sqrt{np(1-p)}\sim N(0,1),$$

这样,就可以利用区间估计的理论了.

大样本下的区间估计可以归结为此类问题.

方法、技巧与典型例题分析

求解本节习题的关键是确定总体X与所讨论的总体比例p,然后按公式进行计算.

例1 某印染厂在配制一种染料时,在40次试验中成功了34次,求配制成功的概率p的置信度为0.95的置信区间.

解 总体是试验的分布,p是成功率.

已知$n=40,\hat{p}=34/40,Z_{0.025}=1.96$,所以$p$的置信度为0.95

的置信区间为

$$\left(\frac{34}{40}-1.96\sqrt{\frac{34}{40}\times\frac{6}{40}\Big/40},\frac{34}{40}+1.96\sqrt{\frac{34}{40}\times\frac{6}{40}\Big/40}\right)$$

$$=(0.7393,0.9607).$$

例 2 某单位工会准备组织一次旅游，为此调查了 100 名职工. 设调查是随机的，且职工人数远远大于 100 人，调查结果显示有 22 人支持，试求支持率 p 的置信度为 0.99 的置信区间.

解 因为 p 的点估计为 $\hat{p}=k/n=0.22$，又

$$\sqrt{\hat{p}(1-\hat{p})/n}=\sqrt{0.22\times0.78/100}=0.0414,$$

而 $Z_{0.01}=2.575$，所以，p 的置信度为 0.99 的置信区间为

$$(0.22\mp0.0414\times2.575)=(0.1134,0.3266).$$

例 3 某纺纱女工看管 800 个纱锭，抽查女工 30 次，每次 1 min，共接头 50 次. 求：

(1) 每分钟断纱次数 $\overline{X}$ 的置信度为 0.95 的置信区间；

(2) 每分钟断纱率 p 的置信度为 0.95 的置信区间.

解 (1) 因为断纱次数 $X\sim\pi(\lambda)$，即 $E(X)=\lambda,D(X)=\lambda$. 而 $\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$，所以有 $E(\overline{X})=\lambda,D(\overline{X})=\lambda/n$.

因为 n 很大，依中心极限定理可以认为，$\overline{X}\sim N(\lambda,\lambda/n)$，所以

$$P\{|\overline{X}-\lambda|/\sqrt{\lambda/n}<Z_{\alpha/2}\}\approx1-\alpha,$$

解得
$$\lambda_{1,2}=\frac{1}{4n}(Z_{\alpha/2}\mp2\sqrt{n\overline{X}+Z_{\alpha/2}^2/4})^2,$$

即 $\overline{X}$ 的置信度为 $1-\alpha$ 的置信区间为 (λ_1,λ_2).

由于 $n=30,n\bar{x}=50,Z_{0.025}=1.96$，代入公式算得 $\lambda_1=1.25,\lambda_2=2.20$，故 $\overline{X}$ 的置信度为 $1-\alpha$ 的置信区间为 $(1.25,2.20)$.

(2) 因为 $\hat{p}$ 很小，$\lambda=n\hat{p}$，所以由题(1)的结果算出

$$\hat{p}_1=\lambda_1/n=1.25/800,\quad \hat{p}_2=\lambda_2/n=2.20/800,$$

于是 p 的置信度为 0.95 的置信区间为

$$(1.25/800, 2.20/800)=(0.00156, 0.00275).$$

例 4 设总体 $X\sim B(1,p)$，其中 p 未知，$0<p<1$，$X_1,X_2,\cdots,X_n$ 为 X 的一个样本. 求 n 很大时，p 的置信度为 $1-\alpha$ 的置信区间.

解 因为 $E(X)=p$，$D(X)=p(1-p)$，n 很大时，依中心极限定理，$\overline{X}\sim N(p,p(1-p)/n)$，即

$$(\overline{X}-p)/\sqrt{p(1-p)/n}\sim N(0,1),$$

所以 p 的置信度为 $1-\alpha$ 的置信区间为

$$\left(\overline{X}-Z_{\alpha/2}\sqrt{\frac{\overline{X}(1-\overline{X})}{n}}, \overline{X}+Z_{\alpha/2}\sqrt{\frac{\overline{X}(1-\overline{X})}{n}}\right).$$

例 5 设总体 X 服从指数分布 $e(\lambda)$，概率密度为

$$f(x)=\begin{cases}\lambda e^{-\lambda x}, & x>0,\\ 0, & \text{其它},\end{cases}$$

其中 $\lambda>0$ 为未知，$x_1,x_2,\cdots,x_n$ 为总体 X 的一个样本(n 很大)，求 λ 的置信度为 $1-\alpha$ 的置信区间.

解 因为 $E(X)=1/\lambda$，$D(X)=1/\lambda^2$，依中心极限定理，当 n 很大时，有

$$\frac{\overline{x}-1/\lambda}{1/(\lambda\sqrt{n})}\overset{\cdot}{\sim}N(0,1),\quad 即\quad \frac{\lambda\overline{x}-1}{1/\sqrt{n}}\overset{\cdot}{\sim}N(0,1),$$

所以
$$P\left\{\left|\frac{\lambda\overline{x}-1}{1/\sqrt{n}}\right|<Z_{\alpha/2}\right\}=1-\alpha.$$

得 λ 的置信度为 $1-\alpha$ 的置信区间为

$$((1-Z_{\alpha/2}/\sqrt{n})/\overline{x}, (1+Z_{\alpha/2}/\sqrt{n})/\overline{x}).$$

例6 根据经验，用船装运玻璃器皿的损坏率不大于5%. 现要估计某船玻璃器皿的损坏率，要求估计与真值不大于5%，在置信度0.90下，应取多大的样本验收？

解 要求估计与真值不大于5%，即要求在置信度0.90下置信区间长度的一半不大于0.05. 因为

$$p=0.05,\quad l=0.10,\quad Z_{0.05}=1.64,$$

所以 $\quad n\geqslant 0.05\times(1-0.05)\times(2\times1.64/0.10)^2=51.1.$

应取样本容量为$n=52$进行验收.

例7 某公司为了推销产品,需要估计有轿车的家庭在当地住户所占的比例,并希望估计误差不要超过5%,要求置信度为0.95.若估计(由经验得出)p不超过0.35,问:应抽多大容量的样本?

解 因为置信区间长度

$$l=2\times0.05=0.1,\quad Z_{0.025}=1.96,\quad p=0.35,$$

所以 $$n\geqslant0.35(1-0.35)(2\times1.96/0.1)^2=349.59,$$

应取样本容量为$n=350$.

若p不能估计,则由$n=(Z_{\alpha/2}/l)^2$,得

$$n\geqslant(1.96/0.1)^2=384.16,$$

可取$n=385$,显然$385>350$.

例8 某市为了决定是否建设高架通道,随机调查了5000个市区居民和2000个郊区居民,分别有2400个和1200个赞成.试在置信度0.90下,求市区和郊区居民赞成人数比例之差的区间估计.

解 赞成人数的点估计为:郊区$\hat{p}_1=1200/2000=0.60$,市区$\hat{p}_2=2400/5000=0.48$. p_1-p_2的点估计$\hat{p}_1-\hat{p}_2=0.12$,所以p_1-p_2的置信度为0.90的置信区间为

$$(0.12\mp Z_{0.05}\sqrt{0.6\times0.4/2000+0.48\times0.52/5000})$$
$$=(0.1084,0.1316).$$

例9 某企业对本单位职工的出勤进行统计,从早班职工中随机抽查了50人,其中全年满勤的有40人;从中班职工中随机抽查了40人,其中全年满勤的有32人.求:在置信度0.95下,早、中班职工满勤比率之差的区间估计.

解 满勤比率的点估计为:早班$\hat{p}_1=40/50=0.8$,中班$\hat{p}_2=35/40=0.875$. p_1-p_2的点估计为

$$\hat{p}_1-\hat{p}_2=0.8-0.875=-0.075.$$

又 $$\sqrt{\hat{p}_1(1-\hat{p}_1)/n_1+\hat{p}_2(1-\hat{p}_2)/n_2}$$

$$=\sqrt{0.8\times0.2/50+0.875\times0.125/40}$$

$$=\sqrt{0.0032+0.0027}=0.077.$$

所以 p_1-p_2 的置信度为0.95的置信区间为

$$(-0.075-Z_{0.025}\times0.077,-0.075+Z_{0.025}\times0.077)$$

$$=(-0.075-1.96\times0.077,-0.075+1.96\times0.077)$$

$$=(0.225,0.076).$$

硕士研究生入学试题分析

一、本章考试要求

1. 理解参数的点估计、估计量与估计值的概念；了解估计量的无偏性、有效性(最小方差性)和相合性(一致性)的概念，并会验证估计量的无偏性.

2. 掌握矩估计法(一阶、二阶矩)和最大似然估计法.

3. 掌握单个正态总体的均值和方差的置信区间的求法.

4. 掌握两个正态总体的均值差和方差比的置信区间的求法.

二、本章重点内容

(一) 点估计

1. 设总体 X 的概率分布为

X	0	1	2	3
p_k	θ^2	$2\theta(1-\theta)$	θ^2	$1-2\theta$

其中 θ ($0<\theta<1/2$) 是未知参数，利用总体 X 的如下样本值：3,1,3,0,3,1,2,3. 求 θ 的矩估计和极大似然估计值. (2002年一)

解 $E(X)=0\times\theta^2+2\theta(1-\theta)+2\theta^2+3(1-2\theta)=3-4\theta$,

$$\bar{x}=\frac{1}{8}(3+1+3+0+3+1+2+3)=2,$$

由矩估计定义，令 $E(X)=\bar{x}$，得 $3-4\theta=2$，所以 θ 的矩估计值

$$\hat{\theta}=1/4.$$

由样本构造似然函数

$$L(\theta)=\theta^2\cdot\theta^2[2\theta(1-\theta)]^2(1-2\theta)^4=4\theta^6(1-\theta)^2(1-2\theta)^4,$$

$$\ln L(\theta)=\ln 4+6\ln\theta+2\ln(1-\theta)+4\ln(1-\theta),$$

$$\frac{\mathrm{d}}{\mathrm{d}\theta}\ln L(\theta)=\frac{6}{\theta}-\frac{2}{1-\theta}-\frac{8}{1-2\theta}=\frac{6-28\theta+24\theta^2}{\theta(1-\theta)(1-2\theta)},$$

令上式等于零,解得

$$\theta_1=\frac{1}{12}(7-\sqrt{13}),\quad \theta_2=\frac{1}{12}(7+\sqrt{13}).$$

$\theta_2>\frac{1}{2}$舍去,于是,θ 的极大似然估计值为

$$\hat{\theta}=\frac{1}{12}(7-\sqrt{13}).$$

2. 设总体 X 的概率密度为

$$f(x)=\begin{cases}6x(\theta-x)/\theta^3, & 0<x<\theta,\\ 0, & 其它,\end{cases}$$

$X_1,X_2,\cdots,X_n$ 是取自总体 X 的简单随机样本.

(1) 求 θ 的矩估计量 $\hat{\theta}$;

(2) 求 $\hat{\theta}$的方差 $D(\hat{\theta})$. (1999 年一)

解 (1) $E(X)=\int_{-\infty}^{+\infty}xf(x)\mathrm{d}x=\int_0^\theta\frac{6x^2(\theta-x)}{\theta^3}\mathrm{d}x=\frac{\theta}{2}$,

记 $\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$,令 $\frac{\theta}{2}=\overline{X}$,得 θ 的矩估计量为 $\hat{\theta}=2\overline{X}$.

(2) $$E(X^2)=\int_0^\theta\frac{6x^3(\theta-x)}{\theta^3}\mathrm{d}x=\frac{6\theta^3}{20},$$

$$D(X)=E(X^2)-[E(X)]^2=\theta^2/20,$$

所以,$\hat{\theta}=2\overline{X}$ 的方差

$$D(\hat{\theta})=D(2\overline{X})=4D(\overline{X})=4\times\frac{1}{n}D(X)=\theta^2/5.$$

3. 设总体 X 的概率密度为

$$f(x)=\begin{cases}(\theta+1)x^{\theta}, & 0<x<1,\\ 0, & \text{其它},\end{cases}$$

其中$\theta>-1$是未知参数，$X_1,X_2,\cdots,X_n$是来自总体X的一个容量为n的简单随机样本，分别用矩估计法和极大似然估计法求θ的估计量. （1997 年一）

解 (1) $E(X)=\int_{-\infty}^{+\infty}(\theta+1)x^{\theta}\cdot x\mathrm{d}x=\dfrac{\theta+1}{\theta+2}$,

令$E(X)=\overline{X}=\dfrac{1}{n}\sum\limits_{i=1}^{n}X_i$，得$\hat{\theta}=\dfrac{2\overline{X}-1}{1-\overline{X}}$，是$\theta$的矩估计量.

(2) 作似然函数

$$L(\theta)=\begin{cases}(\theta+1)^n\left(\prod\limits_{i=1}^{n}x_i\right)^{\theta}, & 0<x_i<1,\\ 0, & \text{其它},\end{cases}$$

取对数，得 $$\ln L(\theta)=n\ln(1+\theta)+\theta\sum_{i=1}^{n}\ln x_i,$$

求导，得 $$\frac{\mathrm{d}}{\mathrm{d}\theta}\ln L(\theta)=\frac{n}{\theta+1}+\sum_{i=1}^{n}\ln x_i,$$

令上式等于零，解得$\hat{\theta}=-1-n\Big/\sum\limits_{i=1}^{n}\ln x_i$，$\hat{\theta}$是$\theta$的极大似然估计量.

4. 设某种元件的使用寿命X的概率密度为

$$f(x,\theta)=\begin{cases}2\theta^{-2(x-\theta)}, & x\geqslant\theta,\\ 0, & \text{其它},\end{cases}$$

其中$\theta>0$为未知参数，又设$x_1,x_2,\cdots,x_n$是X的一组样本观察值，求参数θ的极大似然估计值. （2000 年一）

解 似然函数为

$$L(\theta)=L(x_1,x_2,\cdots,x_n;\theta)$$

$$=\begin{cases}2^n\mathrm{e}^{-2\sum\limits_{i=1}^{n}(x_i-\theta)}, & x_i\geqslant\theta\ (i=1,2,\cdots,n),\\ 0, & \text{其它}.\end{cases}$$

当$x_i\geqslant\theta$ $(i=1,2,\cdots,n)$时，$L(\theta)>0$，取对数，得

$$\ln L(\theta)=n\ln 2-2\sum_{i=1}^{n}(x_i-\theta).$$

因为$\frac{\mathrm{d}}{\mathrm{d}\theta}\ln L(\theta)=2n>0$,所以知$L(\theta)$单调增加.

由于θ要满足$\theta\leqslant x_i\quad(i=1,2,\cdots,n)$,因此当$\theta$取$x_1,x_2,\cdots,x_n$中最小值时,$L(\theta)$取最大值,所以$\theta$的最大似然估计值为

$$\hat{\theta}=\min(x_1,x_2,\cdots,x_n).$$

5. 设总体X的概率密度为

$$p(x,\lambda)=\begin{cases}\lambda a x^{a-1}\mathrm{e}^{-\lambda x^a}, & x>0,\text{其中}\lambda>0\text{未知},\\ 0, & \text{其它},a>0,\text{常数}.\end{cases}$$

据来自总体X的简单随机样本$X_1,X_2,\cdots,X_n$,求λ的极大似然估计量.（1991 年四）

解 似然函数

$$L(\lambda)=\lambda^n\alpha^n\Big(\prod_{i=1}^{n}x_i\Big)^{\alpha-1}\mathrm{e}^{-\lambda\alpha},\ x>0,$$

其中$\lambda>0$未知. 当$x_i>0\ (i=1,2,\cdots,n)$时,$L(\lambda)>0$,取对数,得

$$\ln L(\lambda)=n\ln\lambda+n\ln\alpha+(\alpha-1)\ln\Big(\prod_{i=1}^{n}x_i\Big)+(-\lambda)\sum_{i=1}^{n}x_i^{\alpha},$$

求导,得
$$\frac{\mathrm{d}}{\mathrm{d}\lambda}\ln L(\lambda)=\frac{n}{\lambda}-\sum_{i=1}^{n}x_i^{\alpha},$$

令上式等于零,解得$\hat{\lambda}=n\Big/\Big(\sum_{i=1}^{n}x_i^{\alpha}\Big)$.

6. 设随机变量$X_1,X_2,\cdots,X_n$相互独立且同分布,

$$\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i,\quad S^2=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2,\quad D(X_i)=\sigma^2,$$

则S(　　).（1992 年四）

(A) 是σ的无偏估计;　　(B) 是σ的最大似然估计;

(C) 是σ的一致估计;　　(D) 是与$\overline{X}$相互独立的.

解 选(C). 因为,设$X\sim N(\mu,\sigma^2)$,则

$$B_2=A_2-A_1^2\xrightarrow{p}g(\mu_1,\mu_2)=\mu_2-\mu_1^2$$

$$=E(X^2)-[E(X)]^2=\sigma^2,$$

$$S^2=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2=\frac{n}{n-1}\left[\frac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})^2\right]$$

$$=\frac{n}{n-1}B_2\xrightarrow{p}\sigma^2,$$

即 $S\xrightarrow{p}\sigma$，所以 S 是 σ 的一致估计量.

7. 设总体 X 的概率密度为

$$f(x)=\begin{cases}2\mathrm{e}^{-2(x-\theta)}, & x>\theta,\\ 0, & x\leqslant\theta,\end{cases}$$

其中 $\theta>0$ 是未知参数，从总体中抽取简单随机样本 $X_1,X_2,\cdots,X_n$，记 $\hat{\theta}=\min(X_1,X_2,\cdots,X_n)$.

(1) 求总体 X 的分布函数 $F(x)$；

(2) 求估计量 $\hat{\theta}$ 的分布函数 $F_{\hat{\theta}}(x)$；

(3) 如果用 $\hat{\theta}$ 作为 θ 的估计量，讨论它是否具有无偏性.

(2003 年一)

解 (1) 由定义知

$$F(x)=\int_{\theta}^{x}2\mathrm{e}^{-2(t-\theta)}\mathrm{d}t=\begin{cases}1-\mathrm{e}^{-2(x-\theta)}, & x>\theta,\\ 0, & x\leqslant\theta.\end{cases}$$

(2) $$F_{\hat{\theta}}(x)=P\{\hat{\theta}\leqslant x\}=P\{\min(X_1,X_2,\cdots,X_n)\leqslant x\}$$
$$=1-P\{\min(X_1,X_2,\cdots,X_n)>x\}$$
$$=1-P\{X_1>x\}P\{X_2>x\}\cdots P\{X_n>x\}$$
$$=1-[1-F(x)]^n=\begin{cases}1-\mathrm{e}^{-2n(x-\theta)}, & x>\theta,\\ 0, & x\leqslant\theta.\end{cases}$$

(3) $\hat{\theta}$ 的概率密度为

$$f_{\hat{\theta}}(x)=F'_{\hat{\theta}}(x)=\begin{cases}2n\mathrm{e}^{-2n(x-\theta)}, & x>\theta,\\ 0, & x\leqslant\theta,\end{cases}$$

而 $$E(\hat{\theta})=\int_{\theta}^{\infty}x\cdot 2n\mathrm{e}^{-2n(x-\theta)}\mathrm{d}x=\theta+\frac{1}{2n}\neq\theta,$$

所以，估计量$\hat{\theta}$不具有无偏性.

8. 设总体X的分布函数为

$$F(x;\beta)=\begin{cases}1-1/x^{\beta}, & x>1,\\ 0, & x\leqslant 1,\end{cases}$$

其中未知参数$\beta>1$，$X_1,X_2,\cdots,X_n$为来自总体X的简单随机样本，求：

(1) β的矩估计量；

(2) β的最大似然估计量. (2004年一)

解 X的概率密度为

$$f(x,\beta)=\begin{cases}\beta/x^{\beta+1}, & x>1,\\ 0, & x\leqslant 1.\end{cases}$$

(1) $E(x)=\int_{-\infty}^{+\infty}x\,\dfrac{\beta}{x^{\beta+1}}\mathrm{d}x=\dfrac{\beta}{\beta-1}$. 令$\dfrac{\beta}{\beta-1}=\overline{X}$，得$\beta=\dfrac{\overline{X}}{\overline{X}-1}$，故$\beta$的矩估计量为$\hat{\beta}=\dfrac{\overline{X}}{\overline{X}-1}$.

(2) 作似然函数

$$L(\beta)=\prod_{i=1}^{n}f(x_i,\beta)=\begin{cases}\beta^n/(x_1x_2\cdots x_n)^{\beta+1}, & x_i>1,\\ 0, & \text{其它},\end{cases}$$

当$x_i>1$ $(i=1,2,\cdots,n)$时，有

$$\ln L(\beta)=n\ln\beta-(\beta+1)\sum_{i=1}^{n}\ln x_i.$$

求导，得
$$\frac{\mathrm{d}}{\mathrm{d}\beta}\ln L(\beta)=\frac{n}{\beta}-\sum_{i=1}^{n}\ln x_i,$$

令
$$\frac{\mathrm{d}}{\mathrm{d}\beta}\ln L(\beta)=0,$$

得
$$\beta=n\Big/\sum_{i=1}^{n}\ln x_i.$$

故β的最大似然估计量为$\hat{\beta}=n\Big/\sum_{i=1}^{n}\ln x_i$.

9. 设随机变量X的分布函数为

$$F(x;\alpha,\beta)=\begin{cases}1-(\alpha/x)^{\beta}, & x>\alpha,\\ 0, & x\leqslant\alpha,\end{cases}$$

其中参数 $\alpha>0,\beta>1$. 设 $X_1,X_2,\cdots,X_n$ 为来自总体 X 的简单随机样本，

(1) 当 $\alpha=1$ 时,求未知参数 β 的矩估计量;

(2) 当 $\alpha=1$ 时,求未知参数 β 的最大似然估计量;

(3) 当 $\beta=2$ 时,求未知参数 α 的最大似然估计量.

(2004 年三)

解 设 $x_1,x_2,\cdots,x_n$ 为总体 X 的一组样本值.

(1) 当 $\alpha=1$ 时,X 的概率密度为

$$f(x;\beta)=\begin{cases}\beta/x^{\beta+1}, & x>1,\\ 0, & x\leqslant 1,\end{cases}$$

则
$$E(X)=\int_{-\infty}^{+\infty}x\,\frac{\beta}{x^{\beta+1}}\mathrm{d}x=\frac{\beta}{\beta-1}.$$

令 $\dfrac{\beta}{\beta-1}=\overline{X}$,得 $\beta=\dfrac{\overline{X}}{\overline{X}-1}$,故 β 的矩估计量为 $\hat{\beta}=\dfrac{\overline{X}}{\overline{X}-1}$.

(2) 当 $\alpha=1$ 时,似然函数

$$L(\beta)=\begin{cases}\beta^n/(x_1x_2\cdots x_n)^{\beta+1}, & x_i>1,\\ 0, & x_i\leqslant 1,\end{cases}\quad i=1,2,\cdots,n.$$

当 $x_i>1$ 时,$L(\beta)>0$. 取对数、求导,得

$$[\ln L(\beta)]'=\Big[n\ln\beta-(\beta+1)\sum_{i=1}^{n}\ln x_i\Big]'=n/\beta-\sum_{i=1}^{n}\ln x_i.$$

令 $\dfrac{\mathrm{d}}{\mathrm{d}\beta}\ln L(\beta)=0$,解得 $\beta=n\Big/\sum\limits_{i=1}^{n}\ln x_i$,故 β 的最大似然估计量为

$$\hat{\beta}=n\Big/\sum_{i=1}^{n}\ln x_i.$$

(3) 当 $\beta=2$ 时,X 的概率密度为

$$f(x;\alpha)=\begin{cases}2\alpha^2/x^3, & x>\alpha,\\ 0, & x\leqslant\alpha,\end{cases}$$

其似然函数为

$$L(\alpha)=\begin{cases}2^n\alpha^{2n}/(x_1x_2\cdots x_n)^3, & x_i>\alpha,\\ 0, & x_i\leqslant\alpha,\end{cases}\quad i=1,2,\cdots,n.$$

当$x_i>\alpha$时，α越大，$L(\alpha)$越大，故α的最大似然估计量为$\hat{\alpha}=\min(X_1,X_2,\cdots,X_n)$.

（二）区间估计

1. 设一批零件的长度服从正态分布$N(\mu,\sigma^2)$，其中μ,σ^2均未知. 现从中随机抽取16个零件，测得样本均值$\overline{x}=20$(单位：cm)，样本标准差$S=1$ (cm)，则μ的置信度为0.90的置信区间是(　　).

(A) $\left(20-\frac{1}{4}t_{0.05}(16),20+\frac{1}{4}t_{0.05}(16)\right)$；

(B) $\left(20-\frac{1}{4}t_{0.1}(16),20+\frac{1}{4}t_{0.1}(16)\right)$；

(C) $\left(20-\frac{1}{4}t_{0.05}(15),20+\frac{1}{4}t_{0.05}(15)\right)$；

(D) $\left(20-\frac{1}{4}t_{0.1}(15),20+\frac{1}{4}t_{0.1}(15)\right)$.　　　(2005年三)

解　选(C). 因为按题设，单个正态总体在方差未知时，均值的置信度为α的双侧置信区间为

$$\left(\overline{X}-\frac{S}{\sqrt{n}}t_{\alpha/2}(n-1),\overline{X}+\frac{S}{\sqrt{n}}t_{\alpha/2}(n-1)\right),$$

故选(C).

2. 已知一批零件的长度X(单位：cm)服从正态分布$N(\mu,1)$，从中随机地抽取16个零件，得到长度的平均值为40(单位：cm)，则μ的置信度为0.95的置信区间是________.　　　(2003年一)

解　当σ^2已知时，μ的置信度为$1-\alpha$的置信区间为$(\overline{X}-\sigma/\sqrt{n}Z_{\alpha/2},\overline{X}+\sigma/\sqrt{n}Z_{\alpha/2})$. 将$\sigma^2=1$，$\overline{X}=40$，$n=16$，$\Phi(1.96)=0.975$代入得置信区间为(39.51,40.49).

3. 假设0.50,0.80,1.25,2.00是来自总体X的简单随机样本值，已知$Y=\ln X$服从正态分布$N(\mu,1)$.

(1) 求X的数学期望$E(X)$(记$E(X)$为b)；

(2) 求μ的置信度为0.95的置信区间；

(3) 利用上述结果求 b 的置信度为 0.95 的置信区间.

(2000 年三)

解 (1) Y 的密度函数为

$$f(y)=\frac{1}{\sqrt{2\pi}}e^{-(y-\mu)^2/2},\ -\infty<y<+\infty.$$

作代换 $t=y-\mu$,可得

$$b=E(X)=E(e^Y)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}e^y e^{-(y-\mu)^2/2}dy$$

$$=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}e^{t+\mu}e^{-t^2/2}dt$$

$$=e^{\mu+1/2}\int_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi}}e^{-(t-1)^2/2}dt=e^{\mu+1/2}.$$

(2) 置信度 0.95,即 $\alpha=0.05$,所以 $Z_{0.025}=1.96$. 由于 $\overline{Y}\sim N(\mu,1/4)$,所以 μ 的置信度为 0.95 的置信区间为

$$(\overline{Y}\mp 1.96\times 1/\sqrt{4})=(\overline{Y}-0.98,\overline{Y}+0.98).$$

这里 $\overline{Y}=\frac{1}{4}(\ln 0.5+\ln 0.8+\ln 1.25+\ln 2)=\frac{1}{4}\ln 1=0$,

故参数 μ 的置信度为 0.95 的置信区间为 $(-0.98,0.98)$.

(3) 由 e^x 的严格单调增加,可见 b 的置信度为 0.95 的置信区间为 $(e^{-0.48},e^{1.48})$.

4. 设由来自正态总体 $X\sim N(\mu,0.9^2)$ 的容量为 9 的简单随机样本,得样本均值 $\overline{x}=5$,则未知参数 μ 的置信度为 0.95 的置信区间是________.

(1996 年四)

解 方差 $\sigma^2=0.9^2$,$n=9$,$\overline{x}=5$,所以 μ 的置信度为 0.95 的置信区间为

$$\left(\overline{X}\mp\frac{\sigma}{\sqrt{n}}Z_{\alpha/2}\right)=(5\mp\frac{0.9}{3}Z_{\alpha/2}).$$

查表知,$Z_{0.025}=1.96$,所以置信区间为 $(4.412,5.588)$.

5. 设总体 X 的方差为 1,根据来自 X 的容量为 100 的简单随机样本,测得样本均值为 5,则 X 的数学期望的置信度近似等于 0.95的置信区间为________.

(1993 年四)

解 方差 $\sigma^2=1, n=100, \bar{x}=5, Z_{0.025}=1.96$，所以 μ 的置信度近似等于 0.95 的置信区间为

$$\left(\bar{x}\mp\frac{\sigma}{\sqrt{n}}Z_{\alpha/2}\right)=\left(5\mp\frac{1}{10}\times 1.96\right)=(4.804,5.196).$$

近似是因为总体 X 分布未知，在大样本($n\geqslant 50$)下可以认为近似服从正态分析.

6. 从正态总体 $N(3.4,6^2)$ 抽取容量为 n 的样本，如果要求其样本均值位于区间(1.4,5.4)内的概率不小于0.95，问：样本容量n至少应取多大？ (1998 年一)

解 以 $\overline{X}$ 表示样本均值，则$\dfrac{\overline{X}-3.4}{6/\sqrt{n}}\sim N(0,1)$，从而

$$\begin{aligned}P\{1.4<\overline{X}<5.4\}&=P\{-2<\overline{X}-3.4<2\}\\&=P\{|\overline{X}-3.4|<2\}\\&=P\left\{\frac{|\overline{X}-3.4|}{6/\sqrt{n}}<\frac{2\sqrt{n}}{6}\right\}\\&=2\Phi(\sqrt{n}/2)-1\\&\geqslant 0.95.\end{aligned}$$

表 7.3

Z	$\Phi(Z)$
1.28	0.9
1.645	0.95
1.96	0.975
2.33	0.99

由$\Phi(\sqrt{n}/3)\geqslant 0.975$，查表7.3 知 $\sqrt{n}/3\geqslant 1.96$，于是$n\geqslant(1.96\times 3)^2=34.57$，取 $n=35$.

7. 在天平上反复称量一重量为 a 的物品，假设各次称量的结果相互独立且服从正态分布 $N(a,0.2^2)$. 若以 $\overline{X}_n$ 表示 n 次称量结果的算术平均值，则为使

$$P\{|\overline{X}_n-a|<0.1\}\geqslant 0.95,$$

n 的最小值应不小于自然数________. (1999 年三)

解 这是一个求样本容量的问题. 因为 $0.95=1-\alpha$，所以 $\alpha=0.05$. 而

$$(\overline{X}_n-\alpha)/(\sigma/\sqrt{n})\sim N(0,1),$$

故区间长度为

$$2\times\sigma/\sqrt{n}\times Z_{0.025}\leqslant 0.1\times 2,$$

从而得 $\sqrt{n}\geqslant(2\times0.2\times1.96)/0.2=3.92$，取 $n=16$.

8. 设总体 X 的概率密度为

$$f(x;\theta)=\begin{cases}e^{-(x-\theta)}, & 若\ x\geqslant\theta,\\ 0, & 若\ x<\theta,\end{cases}$$

而 $X_1,X_2,\cdots,X_n$ 是来自总体 X 的简单随机样本，则未知参数 θ 的矩估计量为________.（2002 年卷三）

解 $E(X)=\int_\theta^\infty xe^{-(x-\theta)}dx$

$$=\int_\theta^\infty(x-\theta)e^{-(x-\theta)}dx+\theta\int_\theta^\infty e^{-(x-\theta)}dx(分部)$$

$$=-(\theta+1)e^{-(x-\theta)}\Big|_\theta^\infty=\theta+1,$$

即 $\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i=\theta+1$，所以 θ 的矩估计量为

$$\hat{\theta}=\frac{1}{n}\sum_{i=1}^{n}X_i-1.$$

第八章　假设检验

第一节　正态总体均值的假设检验

主要内容

假设检验问题是统计推断的另一类重要问题.在总体分布函数未知或只知形式不知其参数的情况下,为推断总体的性质而提出某些关于总体的假设.假设检验就是根据样本得到的信息,对提出的假设进行判断:确定接受还是拒绝假设.

一、假设检验的基本概念

1. 假设检验的基本思想

假设检验使用的是概率反证法思想.先对检验对象提出某个假设,然后根据抽样结果,利用小概率原理作出拒绝或接受假设的判断.

2. 假设检验的基本步骤

处理参数的假设检验问题的步骤如下:

(1) 根据实际问题的要求,提出原假设 H_0 和备择假设 H_1;

(2) 给定显著性水平 α 以及样本容量 n;

(3) 确定选用的检验统计量及拒绝域的形式;

(4) 由样本值计算统计量的值;

(5) 按 $P\{拒绝H_0|H_0为真\}=\alpha$,确定拒绝域,并作出判断,确定接受还是拒绝假设.

3. 两类错误

由于是用样本提供的信息来推断总体的特征,而样本的选取

是随机的，所以作出的推断可能出现错误．错误分为两类：

（1）H_0 为真，而作出了拒绝 H_0 的判断，称为犯第一类错误，也称犯“弃真”错误．犯错误的概率记为 α.

（2）H_0 不真，而作出了接受 H_0 的判断，称为犯第二类错误，也称为犯“取伪”错误．犯错误的概率记为 β.

当样本容量确定时，α 与 β 是此消彼长的关系，不可能同时变小．要使 α 和 β 同时变小，只能增加样本的容量．

4. 显著性检验与显著性水平

只控制犯第一类错误的概率，而不考虑犯第二类错误的概率的检验问题，称为显著性检验．

二、正态总体均值的假设检验

1. 单个正态总体均值的假设检验

设 $X_1,X_2,\cdots,X_n$ 是总体 $X\sim N(\mu,\sigma^2)$ 的一个样本．

（1）$\sigma^2=\sigma_0^2$，检验假设 $H_0:\mu=\mu_0$.

选择检验统计量 $U=\dfrac{\overline{X}-\mu}{\sigma/\sqrt{n}}\sim N(0,1)$，拒绝域为

$$|\overline{X}-\mu_0|\geqslant Z_{\alpha/2}\frac{\sigma_0}{\sqrt{n}}.$$

单侧假设检验结果见表8.1.

（2）σ^2 未知，检验假设 $H_0:\mu\neq\mu_0$.

选择检验统计量 $T=\dfrac{\overline{X}-\mu}{s/\sqrt{n}}\sim t(n-1)$，拒绝域为

$$|\overline{X}-\mu_0|\geqslant t_{\alpha/2}(n-1)\frac{S}{\sqrt{n}}.$$

单侧假设检验结果见表8.1.

2. 两个正态总体的假设检验

设 $X_1,X_2,\cdots,X_{n_1}$ 是总体 $X\sim N(\mu_1,\sigma_1^2)$ 的样本，$Y_1,Y_2,\cdots,Y_{n_2}$ 是总体 $Y\sim N(\mu_2,\sigma_2^2)$ 的样本，两者相互独立．它们的样本均值和样本方差分别为 $\overline{X},\overline{Y}$ 和 S_1^2,S_2^2.

表 8.1

条件	假设 H_0	检验统计量	统计量分布	备择假设	拒绝域
已知 $\sigma^2=\sigma_0^2$	$\mu=\mu_0$	$U=\dfrac{X-\mu_0}{\dfrac{\sigma_0}{\sqrt{n}}}$	$N(0,1)$	$\mu\neq\mu_0$	$\lvert u\rvert>Z_{\alpha/2}$
	$\mu\leqslant\mu_0$			$\mu>\mu_0$	$u>Z_\alpha$
	$\mu\geqslant\mu_0$			$\mu<\mu_0$	$u>-Z_\alpha$
σ^2 未知	$\mu=\mu_0$	$T=\dfrac{X-\mu_0}{\dfrac{S}{\sqrt{n}}}$，其中 S^2 为样本方差	$t(n-1)$	$\mu\neq\mu_0$	$\lvert t\rvert<t_{\alpha/2}(n-1)$
	$\mu\leqslant\mu_0$			$\mu>\mu_0$	$t>t_0(n-1)$
	$\mu\geqslant\mu_0$			$\mu<\mu_0$	$t<-t_0(n-1)$

(1) σ_1^2,σ_2^2 已知，检验假设 $H_0:\mu_1=\mu_2$.

用检验统计量 $U=\dfrac{\overline{X}-\overline{Y}}{\sqrt{\sigma_1^2/n_1+\sigma_2^2/n_2}}\sim N(0,1)$，拒绝域为

$$|\overline{X}-\overline{Y}|\geqslant Z_{\alpha/2}\sqrt{\sigma_1^2/n_1+\sigma_2^2/n_2}.$$

单侧假设检验结果见表8.2.

表8.2

条件	假设 H_0	检验统计量	统计量分布	备择假设	拒绝域
已知 σ_1^2,σ_2^2	$\mu_1=\mu_2$	$U=\dfrac{\overline{X}-\overline{Y}}{\sqrt{\dfrac{\sigma_1^2}{n_1}+\dfrac{\sigma_2^2}{n_2}}}$	$N(0,1)$	$\mu_1\neq\mu_2$	$\lvert U\rvert>Z_{\alpha/2}$ $\frac{\alpha}{2}$ $\frac{\alpha}{2}$ $-Z_{\alpha/2}$ O $Z_{\alpha/2}$
	$\mu_1\leqslant\mu_2$			$\mu_1>\mu_2$	$U>Z_\alpha$
	$\mu_1\geqslant\mu_2$			$\mu_1<\mu_2$	$U<-Z_\alpha$
$\sigma_1^2=\sigma_2^2$ 但其值未知	$\mu_1=\mu_2$	$T=\dfrac{\overline{X}-\overline{Y}}{S_W\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}}$，其中 $S_W^2=\dfrac{(n_1-1)S_1^2}{n_1+n_2-2}+\dfrac{(n_2-1)S_2^2}{n_1+n_2-2}$	$t(n_1+n_2-2)$	$\mu_1\neq\mu_2$	$\lvert T\rvert\geqslant t_{\alpha/2}(n_1+n_2-2)$ $\frac{\alpha}{2}$ $\frac{\alpha}{2}$ $-t_{\alpha/2}$ O $t_{\alpha/2}$
	$\mu_1\leqslant\mu_2$			$\mu_1>\mu_2$	$T\geqslant t_\alpha(n_1+n_2-2)$
	$\mu_1\geqslant\mu_2$			$\mu_1<\mu_2$	$T<-t_\alpha(n_1+n_2-2)$

(2) $\sigma_1^2=\sigma_2^2=\sigma^2$，但 σ^2 未知. 检验假设 $H_0:\mu_1=\mu_2$.

选择检验统计量 $T=\dfrac{\overline{X}-\overline{Y}}{S_W\sqrt{1/n_1+1/n_2}}\sim t(n_1+n_2-2)$，拒绝域为

$$|\overline{X}-\overline{Y}|\geqslant t_{\alpha/2}(n_1+n_2-2)S_W\sqrt{1/n_1+1/n_2},$$

其中 $\qquad S_W^2=[(n_1-1)S_1^2+(n_2-1)S_2^2]/(n_1+n_2-2)$.

单侧假设检验结果见表8.2.

疑难解析

1. 什么是显著性检验？其基本思想是什么？

答 只考虑一个假设是否成立的检验称为显著性检验. 其待检假设的一般形式为 $H_0:\theta\in\Theta_0$，其中 Θ_0 为参数空间 Θ 的一个子集. 显著性检验的原则是，只要求犯第一类错误的概率不大于某一正数 α $(0<\alpha<1)$，即若显著性检验法 φ 的拒绝域为 W，则 W 应满足

$$P\{(X_1,X_2,\cdots,X_n)\in W\}\leqslant\alpha,\quad \theta\in\Theta_0.$$

显著性检验法的基本思想是，根据小概率事件在一次试验中一般是不会发生的实际推断原理，依靠从样本得到的信息来判断假设是否可以接受.

对于同一假设，在同一显著性水平下，根据样本 $(X_1,X_2,\cdots,X_n)$，可以建立许多不同的显著性检验. 由于只需满足一个要求 $P\{(X_1,X_2,\cdots,X_n)\in W\}\leqslant\alpha$，所以难以判断诸多显著性检验的优劣.

2. 提出原假设的一般依据是什么？原假设与备择假设在检验假设中的地位是否相同？

答 选择一个问题的哪个结果作原假设，其一般原则如下：

(1) 因为显著性检验只考虑犯第一类错误的概率，所以对犯两类错误可能引起的后果加以比较，将后果严重的列为第一类错误，以 α 来控制它. 如某人去做健康检查，需提出假设“有病”还是“无病”，而把“有病认作无病”的错误显然比把“无病认作有病”的错误后果严重，所以，原假设应取 H_0：有病.

(2) 选择经验的、保守的为原假设. 如某厂一种产品的使用寿命为 μ_0，经过工艺改革，要确认使用寿命是否增加. 这时，取原假设为 $H_0:\mu=\mu_0$.

假设检验控制犯第一类错误的概率，所以检验法是保护原假设，不轻易拒绝原假设的.

3. 显著性检验的反证法与一般的反证法有什么不同?

答 一般的反证法是逻辑上的反证法,即由结果的矛盾而推出假设的错误. 而显著性检验法的反证法是由小概率事件在一次试验中本不该发生,却竟然发生了这一矛盾的结果出发,拒绝原假设 H_0. 由于样本的抽取具有随机性. 因此,由抽样所确定的结果的矛盾也有随机性. 可能出现假设正确而拒绝假设的错误,但犯错误的概率小于α. 所以说,在假设检验中的反证法具有概率的性质,称为“概率反证法”.

4. 为什么在两个总体情形,当待检假设 $H_0:\mu_1=\mu_2$ 时,若 σ_1^2, σ_2^2 未知,要求 $\sigma_1^2=\sigma_2^2=\sigma^2$?

答 当σ_1^2,σ_2^2 未知,且$\sigma_1^2\neq\sigma_2^2$ 时,要检验假设$H_0:\mu_1=\mu_2$ 是不合理的,是不同条件下两总体均值的比较,不能真实反映两总体的某数量指标的优劣. 只有令 $\sigma_1^2=\sigma_2^2$,才能在两总体的相同条件下比较某项数量指标的优劣,作出的判断才有意义.

当σ_1^2 与σ_2^2 是否相等未知时,要先用F 检验法,检验σ_1^2 是否等于σ_2^2. 若相等,再用t 检验法检验两总体的期望 μ_1 和 μ_2 是否相等. 一般,为了使结果更可信,应将α 适当取大一些.

若F 检验的结果是 $\sigma_1^2\neq\sigma_2^2$,这时要检验假设 $H_0:\mu_1=\mu_2$,有以下方法.

(1) 当n_1,n_2 很大时,选用检验统计量

$$U=\frac{\overline{X}-\overline{Y}}{\sqrt{S_1^2/n_1+S_2^2/n_2}}\sim N(0,1),$$

拒绝域为
$$|\overline{X}-\overline{Y}|\geqslant Z_{\alpha/2}\sqrt{S_1^2/n_1+S_2^2/n_2}.$$

(2) 当$n_1=n_2=n$ 时,令 $Z_i=X_i-Y_i$, $i=1,2,\cdots,n$. 记$E(Z_i)=\mu$, $D(Z_i)=\sigma^2$,则 $Z_1,Z_2,\cdots,Z_n$ 为 $N(\mu,\sigma^2)$的样本,$H_0:\mu=0$. 用检验统计量

$$T=\overline{Z}\ \sqrt{n}/S_z^2,\quad S_z^2=\frac{1}{n-1}\sum_{i=1}^{n}(Z_i-\overline{Z})^2,$$

拒绝域为 $$|\overline{Z}|\geqslant t_{\alpha/2}(n-1)\frac{S_z^2}{\sqrt{n}}.$$

(3) 当 $\sigma_1^2=k\sigma_2^2$ 时,用统计量

$$T=\frac{|\overline{X}-\overline{Y}|}{S_W\sqrt{1/n_1+1/(kn_2)}}\sim t(n_1+n_2-2),$$

$$S_W^2=[(n_1-1)S_1^2+(n_2-1)S_2^2]/(n_1+n_2-2),$$

拒绝域为

$$|\overline{X}-\overline{Y}|\geqslant t_{\alpha/2}(n_1+n_2-2)S_W\sqrt{1/n_1+1/(kn_2)}.$$

5. 假设检验与区间估计有何联系与区别?

答 两者选用的检验统计量形式相同.

(1) 假设检验与区间估计的联系.

利用假设检验可以建立区间估计,而利用区间估计也可以得出假设检验.

例如,设 $X\sim N(\mu,\sigma^2)$,$X_1,X_2,\cdots,X_n$ 是 X 的一个样本,方差 σ^2 未知,要检验假设 $H_0:\mu=\mu_0$,H_1,$\mu\neq\mu_0$,则对给定的显著性水平 α,t 检验的接受域为

$$\{|\overline{X}-\mu_0|\leqslant t_{\alpha/2}(n-1)S/\sqrt{n}\},$$

可以写为

$$\overline{X}-t_{\alpha/2}(n-1)S/\sqrt{n}\leqslant\mu_0\leqslant\overline{X}+t_{\alpha/2}(n-1)S/\sqrt{n},$$

将式中 μ_0 改为 μ,则上式恰为 μ 的置信度为 $1-\alpha$ 的区间估计.

反过来,若已得到 μ 的区间估计为

$$\overline{X}-t_{\alpha/2}(n-1)S/\sqrt{n}\leqslant\mu\leqslant\overline{X}+t_{\alpha/2}(n-1)S/\sqrt{n},$$

其置信度为 $1-\alpha$,则只需将式中 μ 改为 μ_0,上式即成为原假设 H_0:$\mu=\mu_0$ 的一个显著性水平为 α 的接受域.

在其它区间估计与假设检验中也存在这种对应关系,但要注意的是,在把假设检验转化为区间估计时,类似公式要存在一个区间形式,否则,不可能实现这种对应.

(2) 假设检验与区间估计的差别.

如果对前例检验假设$\mu=\mu_0$(水平为α)和求μ的置信度为$1-\alpha$的区间估计进行研究,就会发现:

在接受$H_0:\mu=\mu_0$时,所得到的区间估计精度有高有低,当精度较低时,区间的长度较长.显然这时认为$\mu=\mu_0$不够精确.

在拒绝$H_0:\mu=\mu_0$时,也有同样的情形.因为若区间估计的精度很高,虽然区间内不包含μ_0,但区间有可能就在μ_0附近,仍然可以认为$\mu=\mu_0$.

因此通过对具体问题的分析,可以发现区间估计的结论有时可以与假设检验的结论不同.

方法、技巧与典型例题分析

假设检验的方法在主要内容中已详细给出,这里不再赘述.解题的关键在善于分析具体问题的实际情况,确定是单总体还是双总体,是均值检验还是方差检验,原假设是什么.然后选择适当的统计量,依据样本值计算统计量的值,与根据显著性水平查分布表得到的数值相比较,作出接受还是拒绝原假设H_0的判断.

例1 设总体$X\sim N(\mu,1)$,$x_1,x_2,\cdots,x_{10}$是X的一组样本观察值,要在$\alpha=0.05$的水平下检验假设$H_0:\mu=0,H_1:\mu\neq 0$.拒绝域为$R=\{|\overline{x}>c|\}$.

(1) 求c的值;(2)若已知$\overline{x}=1$,是否可据此样本推断$\mu=0$;(3)若以$R=\{|\overline{x}|\geqslant 1.15\}$作为检验$H_0:\mu=0$的拒绝域,求试验的显著性水平$\alpha$.

解 (1) 是单总体下均值μ的双侧检验,检验统计量是

$$U=\frac{\overline{X}-\mu}{\sigma_0/\sqrt{n}}\sim N(0,1).$$

对$\alpha=0.05$,查正态分布表知$P\{|u|\geqslant 1.96\}=0.05$,从而得拒绝域为$R=\{|u|\geqslant 1.96\}=\{|\sqrt{10}\overline{x}|\geqslant 1.96\}=\{|\overline{x}|\geqslant 0.62\}$.于是知,$c=0.62$.

(2) 由 $\bar{x}=1>0.62$ 即 $\bar{x}\in R$ 知，不能由样本推断 $\mu=0$ 成立.

(3)
$$
\begin{aligned}
P\{|\bar{x}|\geqslant 1.15\} &= P\{\sqrt{10}\bar{X}\geqslant 1.15\sqrt{10}\} \\
&= 1-P\{\sqrt{10}\bar{X}\leqslant 1.15\sqrt{10}\} \\
&= 1-[2\Phi(3.64)-1]=0.0003.
\end{aligned}
$$

而显著性水平即为 $\mu=0$ 成立时拒绝 $H_0:\mu=0$ 的概率，所以 $\alpha=0.0003$.

例 2 某种零件的长度服从正态分布，方差 $\sigma^2=1.21$，现从零件堆中随机抽取 6 件，测得长度(单位:mm)为

32.46，31.54，30.10，29.76，31.67，31.23.

问:当显著性水平为 $\alpha=0.01$ 时，能否认为这批零件的平均长度为 32.50 mm?

解 是 $\sigma^2=1.21$ 已知、单总体下均值 μ 的双侧检验. 待检假设 $H_0:\mu=32.50$，$\alpha=0.01$.

因为算得 $\bar{x}=31.13$， $\sigma=1.1$，

所以用检验统计量 U，得

$$
|u|=\frac{|31.13-32.50|}{1\times 1/\sqrt{6}}=3.05,
$$

查表知 $Z_{0.005}=2.58$，经比较知 $|u|=3.05>Z_{0.005}=2.58$，所以拒绝 H_0，认为该批零件的平均长度不是 32.50 mm.

注意:在解题过程中，拒绝域的两种形式

$$
|\bar{X}-\mu|\geqslant Z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \quad \text{和} \quad \frac{|\bar{X}-\mu|}{\sigma/\sqrt{n}}\geqslant Z_{\alpha/2}
$$

是一致的，用哪个都可以. 在其它检验中同样如此.

例 3 某厂生产的一种产品的质量指标为 $X\sim N(12,1)$. 改革加工工艺后，从新生产的产品中随机抽取了 100 件，测得 $\bar{x}=12.5$. 设方差没有改变，问:改革工艺后该产品的质量指标是否有明显变化($\alpha=0.10$)?

解 是 $\sigma^2=1$ 已知、单总体下均值 μ 的双侧检验. 待检假设

$$
H_0:\mu=12, \quad \alpha=0.10.
$$

因为 $\bar{x}=12.5$， $n=100$， $\sigma=1$，所以用检验统计量U，得

$$|u|=\frac{12.5-12}{1/\sqrt{100}}=5,$$

查正态分布表知$Z_{0.05}=1.64$，经比较知$|u|=5>Z_{0.05}=1.64$，所以拒绝H_0，即认为新工艺下产品的质量指标有明显变化.

例4 某厂生产的电视机显像管的使用寿命(单位:h)为X，$X\sim N(5000,90000)$.使用新设备后要了解使用寿命是否有提高，抽取了36只显像管进行测试.以$H_0:\mu=5000$为原假设，求检验法的拒绝域与接受域(规定:以$\bar{x}>5100$为显像管寿命有提高，$\bar{x}\leqslant 5100$为显像管寿命没有提高)，并求犯第一类错误的概率.

解 总体$X\sim N(\mu,90000)$，μ未知，方差不变，待检假设

$$H_0:\mu=5000,\quad H_1:\mu>5100.$$

$X_1,X_2,\cdots,X_{36}$为X的一个样本.拒绝域为

$$R=\left\{\frac{1}{36}\sum_{i=1}^{36}X_i>5100\right\},$$

接受域为
$$R=\left\{\frac{1}{36}\sum_{i=1}^{36}X_i\leqslant 5100\right\}.$$

因为
$$\bar{X}=\frac{1}{36}\sum_{i=1}^{36}X_i\sim N(\mu,2500),$$

所以，此检验法犯第一类错误的概率为

$$\begin{aligned}\alpha&=P\{\text{拒绝}H_0|H_0\text{为真}\}=P\left\{\frac{1}{36}\sum_{i=1}^{36}X_i>5100|\mu=5000\right\}\\&=\int_{5100}^{+\infty}\frac{1}{\sqrt{2\pi}\sqrt{2500}}e^{-(x-5000)^2/5000}dx=\int_2^{\infty}\frac{1}{\sqrt{2\pi}}e^{-u^2/2}du\\&=1-\Phi(2)=1-0.9772=0.0228.\end{aligned}$$

例5 某种电器零件的平均电阻为2.64 Ω，改变工艺后，测得100个零件的平均电阻为2.62 Ω.设改变工艺前后的电阻的方差保持在0.06^2，问:新工艺对此零件的电阻有无显著的影响($\alpha=0.01$)?

解 是$\sigma^2=0.06^2$已知、单总体下均值μ的双侧检验.待检假

设 $H_0: \mu=\mu_0=2.64, \alpha=0.01$.

因为 $\bar{x}=2.62, \quad n=100$,

所以用检验统计量 U，得

$$|u|=\frac{|\bar{x}-\mu|}{\sigma/\sqrt{n}}=\frac{|2.62-2.64|}{0.06/10}=3.333.$$

而 $Z_{\alpha/2}=Z_{0.005}=2.576$，经比较知 $|u|=3.333>Z_{0.005}=2.576$，故拒绝 H_0，认为改变工艺后，电阻零件有显著变化.

例 6 设某次考试的成绩服从正态分布. 随机抽取了 36 位考生的成绩，算得平均分为66.5 分，标准差$S=15$. 问：在显著性水平 $\alpha=0.05$ 下，是否可以认为这次考试的平均成绩为 70 分？

解 是σ^2 未知、单总体下均值 μ 的双侧检验. 待检假设

$$H_0: \mu=\mu_0=70, \quad \alpha=0.05.$$

因为 $\bar{x}=66.5, S=15, n=36$，所以用检验统计量 T，得

$$|t|=\frac{|66.5-70|}{15/\sqrt{36}}=1.4.$$

查表知 $t_{0.025}(36-1)=2.0301$，经比较知 $|t|=1.4<t_{0.025}(35)=2.0301$，故接受 H_0，认为这次考试的平均成绩为 70 分.

例7 设番茄汁罐头中VC(维生素C)含量服从正态分布. 按照规定，每盒罐头汁中 VC 的平均含量不得少于 21 mg. 现从一批罐头中随机抽取了 16 盒，算得 $\bar{x}=23. S^2=3.9^2$，问：这批罐头的 VC 含量是否合格($\alpha=0.05$)？

解 是σ^2 未知、单总体下均值 μ 的单侧检验，待检假设 $H_0: \mu \leqslant 21$，备择假设 $H_1: \mu>21, \alpha=0.05$.

因为 $\bar{x}=23, S=3.9, n=16$，所以用检验统计量 T，得

$$t=\frac{23-21}{3.9/\sqrt{16}}=\frac{2}{3.9}\times 4=2.0513.$$

查表知 $t_{0.05}(15)=1.7531$，经比较知 $t=2.0513>t_{0.05}(15)=1.7531$，故拒绝 H_0，认为这批罐头的 VC 含量合格.

例8 某炼铁厂的铁液含碳量(质量分数，%)在正常状态下服

从 $N(4.55,0.11^2)$，当日随机测得5炉铁液的含碳量为

$$4.28,\ 4.40,\ 4.42,\ 4.35,\ 4.37,$$

问：在方差不变的假设下，铁液含碳量的均值是否显著降低（$\alpha=0.05$）？

解 是 σ^2 已知、单总体下关于均值 μ 的单侧检验. 待检假设 $H_0:\mu=4.55$，备择假设 $H_1:\mu<4.55$，$\alpha=0.05$.

因为算得 $\bar{x}=4.364$，$n=5$，$\sigma=0.11$，所以用检验统计量 U，得

$$u=\frac{4.364-4.55}{0.11/\sqrt{5}}=-3.781.$$

查表知 $Z_{0.95}=-1.645$，经比较知 $u=-3.781<Z_{0.05}=-1.645$，故拒绝 H_0，认为当日铁液含碳量的均值有显著降低.

例9 为测定某种药物对人的血压有无疗效，测定了10名试验者在服药前后的血压，得血压差值的数据为

$$6,\ 8,\ 4,\ 6,\ -3,\ 7,\ 2,\ 6,\ -2,\ -1,$$

问：在 $\alpha=0.05$ 下，能否认为该药物能够改变人的血压？

解 是 σ^2 未知、单总体下均值 μ 的双侧检验. 待检假设

$$H_0:\mu=0,\quad \alpha=0.05.$$

因为算得 $\bar{x}=3.3$，$S^2=7.014$，$n=10$，

所以用检验统计量 T，得

$$|t|=\frac{3.3-0}{\sqrt{7.014}}\times\sqrt{10}=3.9409.$$

查表知 $t_{0.025}(9)=2.2622$，经比较知 $|t|=3.9409>t_{0.025}(9)=2.2622$，故拒绝 H_0，认为该药物能够改变人的血压.

例10 一种燃料的辛烷等级服从正态分布，其平均等级为98，标准差为0.8. 今从一批新燃料中随机抽取25桶，算得样本均值为97.7. 假定标准差与原来一样，问：新燃料油的辛烷平均等级是否比原燃料辛烷平均等级偏低（$\alpha=0.05$）？

解 是 σ^2 已知、单总体下均值 μ 的单侧检验. 待检假设 $H_0:\mu\geqslant98$，备择假设 $H_1:\mu<98$.

因为$\bar{x}=97.7$，$\sigma=0.8$，$n=25$，所以用检验统计量U，得

$$u=\frac{\bar{x}-\mu}{\sigma/\sqrt{n}}=\frac{-0.3}{0.8}\times 5=-1.875.$$

查表知$Z_{0.05}=-1.645$，经比较知$u=-1.875<Z_{0.95}=-1.645$，故拒绝H_0，认为新燃料油的辛烷平均等级要比原燃料辛烷平均等级偏低.

例 11 某化工厂生产的一种产品的含硫量(质量分数，%)在正常情况下服从正态分布$N(4.55,\sigma^2)$. 为了解设备维修后产品含硫量μ是否改变，测试了5个产品，测得它们含硫量为

$$4.28,\ 4.40,\ 4.42,\ 4.35,\ 4.37,$$

试在下列两种情形下分别检验

$$H_0:\mu=4.55,\quad H_1:\mu\neq 4.55.$$

假定方差不变，$\alpha=0.05$.

(1) $\sigma^2=0.01$；　　(2) σ^2 未知.

解 因为算得$\bar{x}=4.364$，$S^2=0.00293$，所以：

(1) $\sigma^2=0.01$ 已知，应选用检验统计量U，于是

$$|u|=\frac{0.186}{0.1}\times\sqrt{5}=4.16.$$

查表知$Z_{0.025}=1.96$，经比较知$|u|=4.16>Z_{0.025}=1.96$，故拒绝H_0，认为含硫量发生了变化.

(2) σ^2 未知，应选用检验统计量T，于是

$$|t|=\frac{0.186}{\sqrt{0.00293}}\times\sqrt{5}=7.6835.$$

查表知$t_{0.025}(4)=2.7764$，经比较知$|t|=7.6835>t_{0.025}(4)=2.7764$，故拒绝$H_0$，认为含硫量发生了变化.

例 12 某种产品在处理前后的含脂率(质量分数，%)分别为X和Y，$X\sim N(\mu_1,\sigma^2)$，$Y\sim N(\mu_2,\sigma^2)$. 从产品中任取10件，测得它们在处理前后的含脂率如表8.3所示. 试在显著性水平$\alpha=0.05$下，检验处理前后的含脂率有无显著变化.

表 8.3

处理前	0.19	0.18	0.21	0.30	0.66	0.42	0.08	0.12	0.30	0.27
处理后	0.19	0.24	1.04	0.08	0.20	0.12	0.31	0.29	0.13	0.07

解 是$\sigma_1^2=\sigma_2^2=\sigma^2$但$\sigma^2$未知、两个正态总体下均值的假设检验，待检假设$H_0:\mu_1=\mu_2$，是$\alpha=0.05$下的双侧检验.

因为 $\bar{x}=0.273$， $\bar{y}=0.267$，

$$S_1=0.1677,\quad S_2=0.2839,\quad n_1=n_2=10,$$

所以由检验统计量

$$T=\frac{\overline{X}-\overline{Y}}{S_W\sqrt{1/n_1+1/n_2}}\sim t(n_1+n_2-2)$$

得
$$|t|=\frac{0.273-0.267}{0.233\times 0.4714}=0.055.$$

查表知$t_{0.025}(18)=2.101$，经比较知$|t|=0.055<t_{0.025}(18)=2.101$，故接受$H_0$，认为处理前后的产品含脂率无显著变化.

例 13 某林场采用两种方案作杨树育苗试验.已知两种方案下苗高均服从正态分布，标准差分别为$\sigma_1=20$，$\sigma_2=18$.现各抽60棵树苗作样本，测得苗高$\bar{x}_1=59.34$ cm，$\bar{x}_2=49.16$ cm，试以95%的可靠性估计两种方案对杨树苗的高度有无影响.

解 是σ_1^2，σ_2^2已知、双总体下均值的假设检验，待检假设$H_0:\mu_1=\mu_2$，是$\alpha=0.05$下的双侧检验.

因为$\bar{x}_1=59.34$，$\bar{x}_2=49.16$，$\sigma_1=20$，$\sigma_2=18$，$n_1=n_2=60$，所以由检验统计量U，得

$$|u|=\frac{59.34-49.16}{\sqrt{20^2+18^2}}\times\sqrt{60}=2.93.$$

查表知$Z_{0.025}=1.96$，经比较知$|u|=2.93>Z_{0.025}-1.96$，故拒绝H_0，认为两种方案对杨树苗的高度有显著影响.

例14 甲、乙两台机床加工同一种产品，设两台机床加工的零件外径都服从正态分布，标准差分别为$\sigma_1=0.20$，$\sigma_2=0.40$.现从

加工的零件中分别抽取 8 件和 7 件，测得其外径（单位：cm）如表 8.4所示. 试在 $\alpha=0.05$ 下检验两机床加工的零件外径有无显著的差异.

表 8.4

甲	20.5	19.8	19.7	20.4	20.1	20.0	19.0	19.9
乙	19.7	20.8	20.5	19.8	19.4	20.6	19.2	

解 是 σ_1^2,σ_2^2 已知、双总体下均值的假设检验. 待检假设 $H_0:\mu_1=\mu_2$，是 $\alpha=0.05$ 下的双侧检验.

已知 $n_1=8,n_2=7,\sigma_1=0.20,\sigma_2=0.40$，算得 $\bar{x}=19.93,\bar{y}=20$，所以由检验统计量

$$U=\frac{|\overline{X}-\overline{Y}|}{\sqrt{\sigma_1^2/n_1+\sigma_2^2/n_2}}$$

得
$$|u|=\frac{0.07}{\sqrt{0.04/8+0.16/7}}=0.42.$$

查表知 $Z_{0.025}=1.96$，经比较知 $|u|=0.42<Z_{0.025}=1.96$，故接受 H_0，认为两机床加工的零件外径无显著的差异.

例 15 在酿造啤酒中要形成致癌物质 NDMA，现测得旧、新两种工艺过程中形成的 NDMA 含量（质量分数，$\times10^{-9}$）如表 8.5所示. 设两样本都服从正态分布，且总体方差相等，作检验

$$H_0:\mu_1-\mu_2\leqslant 2,\quad H_1:\mu_1-\mu_2>2\ (\alpha=0.05).$$

表 8.5

旧过程	6	4	5	5	6	5	5	6	4	6	7	4
新过程	2	1	2	2	1	0	3	2	1	0	1	3

解 是 $\sigma_1^2=\sigma_2^2=\sigma^2$ 但 σ^2 未知、双总体下均值的假设检验，待检假设 $H_0:\mu_1-\mu_2\leqslant 2$，是 $\alpha=0.05$ 下的单侧检验.

因为算得

$$n_1=n_2=12,\quad \bar{x}=5.25,\quad \bar{y}=1.5,\quad S_1=0.965,\quad S_2=1,$$

用检验统计量

$$T=\frac{\bar{x}-\bar{y}-\delta}{S_W\sqrt{1/n_1+1/n_2}}$$

得
$$t=\frac{5.25-1.5-2}{0.983\times 0.408}=4.3633,$$

此处
$$S_W^2=\frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2},\quad \delta=\mu_1-\mu_2.$$

查表知 $t_{0.05}(22)=1.7171$，经比较知 $t=4.3633>t_{0.05}(22)=1.7171$，故拒绝 H_0，认为 $\mu_1-\mu_2>2$.

例16 据推测认为，矮个子的人比高个子的人寿命要长一些. 下面将美国 31 个自然死亡的总统分为矮个子与高个子两类（以 172.72 cm（68 in）为界），其寿命如表8.6 所示. 设两个寿命总体均服从正态分布，且方差相等. 问：数据显示是否符合推测（$\alpha=0.05$）？

表 8.6

矮个子	85	79	67	90	80				
高个子	68	53	63	70	88	74	64	66	60
	60	78	71	67	90	73	71	77	72
	57	78	67	56	63	64	83	65	

解 是 $\sigma_1^2=\sigma_2^2=\sigma^2$ 但 σ^2 未知、双总体下均值的假设检验，待检假设 $H_0:\mu_1\leqslant\mu_2, H_1:\mu_1>\mu_2$，是 $\alpha=0.05$ 下的单侧检验.

因为 $\bar{x}=80.2, \bar{y}=69.15, S_1=8.585, S_2=9.315, n_1=5, n_2=26$. 由检验统计量

$$T=\frac{\bar{X}-\bar{Y}}{S_W\sqrt{1/n_1+1/n_2}}$$

得
$$t=\frac{80.2-69.15}{9.218\times 0.488}=2.456.$$

查表知 $t_{0.05}(29)=1.6991$，经比较知 $t=2.456>t_{0.05}(29)=1.6991$，故拒绝 H_0，认为推测正确，矮个子人的寿命高于高个子人的寿命.

例 17　某试验室分别在70 ℃和80 ℃的温度下对某项指标分别作了8次重复试验，测得该项指标的数据如表8.7所示. 由经验知数据服从正态分布，且方差相等. 问：在$\alpha=0.05$时，是否可以认为数学期望也相等？

表 8.7

70 ℃	20.5	18.8	19.8	20.9	21.2	21.0	19.5	21.5
80 ℃	20.3	18.8	20.0	20.1	20.2	19.1	19.0	17.7

解　是$\sigma_1^2=\sigma_2^2=\sigma^2$但$\sigma^2$未知、双总体下均值$\mu$的假设检验，待检假设$H_0:\mu_1=\mu_2$，是$\alpha=0.05$下的双侧检验.

因为

$n_1=n_2=8$，　$\bar{x}=20.4$，　$S_1^2=0.8857$，　$\bar{y}=19.4$，　$S_2^2=0.8286$.

由检验统计量T，得

$$|t|=\frac{20.4-19.4}{0.9258\times 0.5}=2.1603.$$

查表知$t_{0.025}(14)=2.1448$，经比较知$|t|=2.1603>t_{0.025}(14)=2.1448$，故拒绝$H_0$，认为数学期望不相等. 在实际问题中，如果两个数据过于接近，则不宜作出结论，应重新抽样试验.

例 18　从总体$X\sim N(\mu_1,\sigma^2)$，$Y\sim N(\mu_2,\sigma^2)$中分别抽取一个容量为n的样本，两样本相互独立，试设计一种较简单的检验法，检验假设$H_0:\mu_1=\mu_2$　（显著性水平α）.

解　可以设$\mu=\mu_1-\mu_2$，则样本值为

$$d_i=x_i-y_i,\ i=1,2,\cdots,n.$$

总体$Z=X-Y\sim N(\mu,2\sigma^2)$，待检假设$H_0:\mu=0$.

设$\bar{d}$和S_d为样本均值与样本标准差，因$2\sigma^2$未知，用检验统计量T，于是

$$T=\frac{\bar{d}}{S_d/\sqrt{n}}\sim t(n-1),$$

拒绝域为　　$|t|>t_{\alpha/2}(n-1)$.

例 19　从总体$X\sim N(\mu_1,\sigma^2)$取样本$X_1,X_2,\cdots,X_{n_1}$，从总体

$Y \sim N(\mu_2, \sigma^2)$中取样本$Y_1, Y_2, \cdots, Y_{n_2}$，两样本相互独立. 待检假设$H_0: \mu_1 = k\mu_2$ $(k \neq 0)$，试确定检验统计量与拒绝域.

解 因为 $\bar{x} \sim N\left(\mu_1, \frac{\sigma^2}{n_1}\right)$， $\bar{y} \sim N\left(\mu_2, \frac{\sigma^2}{n_2}\right)$，

所以
$$\bar{x} - k\bar{y} \sim N\left(\mu_1 - k\mu_2, \frac{\sigma^2}{n_1} + \frac{k^2\sigma^2}{n_2}\right),$$

于是
$$\frac{(\bar{x} - k\bar{y}) - (\mu_1 - k\mu_2)}{\sigma}\sqrt{\frac{n_1 n_2}{k^2 n_1 + n_2}} \sim N(0,1).$$

又
$$\frac{(n_1-1)S_1^2}{\sigma^2} \sim \chi^2(n_1-1), \frac{(n_2-1)S_2^2}{\sigma^2} \sim \chi(n_2-1),$$

所以在S_1^2与S_2^2独立时，有
$$\frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{\sigma^2} \sim \chi^2(n_1+n_2-2).$$

给出检验统计量(由t分布定义)
$$T = \frac{\overline{X} - k\overline{Y}}{S_W}\sqrt{\frac{n_1 n_2}{k^2 n_1 + n_2}} \sim t(n_1+n_2-2),$$

拒绝域为
$$|t| > t_{\alpha/2}(n_1+n_2-2).$$

这是因为，若H_0成立，$\mu_1 - k\mu_2 = 0$，而
$$S_W = \sqrt{[(n_1-1)S_1^2 + (n_2-1)S_2^2]/(n_1+n_2-2)}.$$

例20 某厂A、B两个化验室每天同时从工厂的冷却水中取样，以测定水中的含氯量(质量分数，$\times 10^{-6}$)，记录7 d的数量如表8.8所示. 设$d_i = x_i - y_i$，$i = 1, 2, \cdots, 7$，来自正态总体，问：两化验室的测定结果有无显著差异($\alpha = 0.001$)?

表8.8

A室	1.15	1.86	0.75	1.82	1.14	1.65	1.90
B室	1.00	1.90	0.90	1.80	1.20	1.70	1.95

解 利用例18的方法. 待检假设$H_0: \mu = 0$，d_i为

0.15，−0.04，−0.15，0.02，−0.06，−0.05，−0.05，

算得 $\bar{d} = -0.026$， $S_d = 0.092$， $n = 7$，

所以由
$$T=\frac{\overline{d}}{S_d/\sqrt{n}}$$
得
$$|t|=\frac{0.026}{0.092}\times\sqrt{7}=0.7477.$$
查表知 $t_{0.005}(6)=3.7074$，经比较知 $|t|=0.7477<t_{0.005}(6)=3.7074$，故接受 H_0，认为两化验室的测定结果无显著差异.

例18与例20也称成对数据的检验，其特征是 d_i 来自正态总体. 用于比较两种方法、两种产品或两种仪器的差异，常在相同条件下(如方差相等)作对比试验，取得成对的观察值，然后分析数据作出推断.

例 21 为确定某工艺对降低橡胶制品中含硫量(质量分数,%)，在产品中随机抽取了10件样品，记录了处理前后含硫量如表8.9所示. 试根据数据确定这种工艺对降低橡胶制品中含硫量的变化有无作用($\alpha=0.05$).

表 8.9

处理前	6.05	5.75	7.12	7.10	6.80	6.55	5.90	7.24	5.75	7.30
后处理	5.68	5.40	5.90	6.05	6.00	5.55	5.15	6.34	5.60	6.40

解 设 $d_i=x_i-y_i$, $i=1,2,\cdots,10$, d_i 来自正态总体，为

0.37,0.35,1.22,,1.05,0.80,1.00,0.75,0.90,0.15,0.90,

待检假设 $H_0:\mu\geqslant0$, $H_1:\mu<0$，是单侧检验.

因为 $\overline{d}=0.749$, $S_d=0.3489$, $n=10$,

所以由
$$T=\frac{\overline{d}}{S_d/\sqrt{n}}$$
得
$$t=\frac{0.749}{0.3489}\times\sqrt{10}=6.7887.$$
查表知 $t_{0.05}(9)=1.8331$，经比较知 $t=6.7887>t_{0.05}(9)=1.8331$，故拒绝 H_0，认为工艺对降低橡胶制品含硫量有显著作用.

原假设的选择通常基于这样一种想法：若我们希望某项结论出现时，就故意把该结论不出现作为原假设. 这样，如果能在 α 很

小时拒绝原假设，则结论出现就得到有力的支持. 反之，若以结论出现为原假设，则即被接受也只能说明假设与试验数据相容，不能反映受到数据的有力支持.

例 22 某药厂生产一种新止痛片，厂方希望验证服用新药片后至开始起作用的时间间隔较原有止痛片至少缩短一半，故需检验假设

$$H_0:\mu_1=2\mu_2,\quad H_1:\mu_1>2\mu_2.$$

这里 μ_1 和 μ_2 是两种止痛片起作用时间间隔的总体均值. 设两总体都是正态总体，σ_1^2,σ_2^2 为已知，$x_1,x_2,\cdots,x_{n_1}$ 和 $y_1,y_2,\cdots,y_{n_2}$ 为取自两总体的独立样本. 试在显著性水平 α 下给出 H_0 的拒绝域.

解 本例属于例 19 的情形，$k=2$，但 σ_1^2,σ_2^2 已知.

设 $\quad \bar{x}\sim N(\mu_1,\sigma_1^2/n_1),\quad \bar{y}\sim N(\mu_2,\sigma_2^2/n_2),$

则 $\quad \bar{x}-2\bar{y}\sim N(\mu_1-\mu_2,\sigma_1^2/n_1+\sigma_2^2/n_2).$

当 H_0 为真时，$\quad \dfrac{\bar{x}-2\bar{y}}{\sqrt{\sigma_1^2/n_1+4\sigma_2^2/n_2}}\sim N(0,1).$

因为是单侧检验，故拒绝域为

$$\frac{\bar{x}-2\bar{y}}{\sqrt{\sigma_1^2/n_1+4\sigma_2^2/n_2}}\geqslant Z_\alpha.$$

对 $\sigma_1^2\neq\sigma_2^2$ 但大样本的情形，可以认为总体近似正态分布，且以 S_1^2,S_2^2 代替 σ_1^2,σ_2^2，用检验统计量 U 检验假设.

例 23 某细纱车间在两种工艺条件下各抽取 100 个试样，测得细纱强力的数据，经计算得

甲工艺：$n_1=100,\quad \bar{x}=280,\quad S_1=28;$

乙工艺：$n_2=100,\quad \bar{y}=286,\quad S_2=28.5.$

问：在 $\alpha=0.05$ 下，两种工艺条件对细纱强力有无显著影响？

解 $\sigma_1^2\neq\sigma_2^2$ 且未知，$n_1=n_2=100$，待检假设 $H_0:\mu_1=\mu_2$. 由检验统计量 U，得

$$|u|=\frac{|280-286|}{\sqrt{28^2/100+28.5^2/100}}=\frac{6}{3.995}=1.50.$$

查表知 $Z_{0.025}=1.96$，经比较知 $|u|=1.50<Z_{0.025}=1.96$，故接受 H_0，认为两种工艺条件对细纱强力无显著影响.

例24 甲、乙两机床加工同一种零件，抽样测量其产品的尺寸(单位：mm)，经计算得

甲机床：$n_1=80$， $\overline{x}=33.75$， $S_1=0.1$；

乙机床：$n_2=100$， $\overline{y}=34.15$， $S_2=0.15$.

问：在 $\alpha=0.01$ 下，两机床加工的产品尺寸有无显著差异？

解 $n\geqslant 50$ 时，即可认为是大样本问题. σ_1^2,σ_2^2 均未知，待检假设 $H_0:\mu_1=\mu_2$. 由检验统计量 U，得

$$|u|=\frac{|33.75-34.15|}{\sqrt{0.1^2/80+0.15^2/80}}=\frac{0.4}{0.02}=20.00.$$

查表知 $Z_{0.005}=2.57$，经比较知 $|u|=20.00>Z_{0.005}=2.57$，故拒绝 H_0，认为两机床加工的产品尺寸有显著差异.

在更多的时候，我们只得到两个样本的一批数据，不知 $\mu_1,\mu_2,\sigma_1^2,\sigma_2^2$，也不知 σ_1^2 与 σ_2^2 是否相等. 这时，要检验均值 μ，就要先检验方差齐性，即 $\sigma_1^2=\sigma_2^2$ 是否成立. 若成立，则可用检验统计量 T 来检验均值. 这类例题，将在下节中演绎.

第二节 正态总体方差的假设检验

主要内容

1. 单个正态总体方差的检验

设 $X_1,X_2,\cdots,X_n$ 为来自总体 $X\sim N(\mu,\sigma^2)$ 的一个样本.

(1) μ 未知时，待检假设 $H_0:\sigma^2=\sigma_0^2$.

用检验统计量 $\chi^2=\dfrac{(n-1)S^2}{\sigma_0^2}\sim\chi^2(n-1)$，拒绝域为

$$\chi^2 \leqslant \chi^2_{1-\alpha/2}(n-1) \quad 或 \quad \chi^2 \geqslant \chi^2_{\alpha/2}(n-1).$$

(2) μ 已知时，待检假设 $H_0: \sigma^2 = \sigma_0^2$.

用检验统计量 $\chi^2 = \dfrac{\sum\limits_{i=1}^{n}(X_i-\mu)^2}{\sigma_0^2} \sim \chi^2(n)$，拒绝域为

$$\chi^2 \leqslant \chi^2_{1-\alpha/2}(n) \quad 或 \quad \chi^2 \geqslant \chi^2_{\alpha/2}(n).$$

单侧检验如表 8.10 所示.

表 8.10

条件	假设 H_0	检验统计量	统计量分布	备择假设	拒绝域
μ 未知	$\sigma^2 = \sigma_0^2$	$\chi^2 = \dfrac{(n-1)S^2}{\sigma_0^2}$	$\chi^2(n-1)$	$\sigma^2 \neq \sigma_0^2$	$\chi^2 \geqslant \chi^2_{\alpha/2}(n-1)$ 或 $\chi^2 \leqslant \chi^2_{1-\alpha/2}(n-1)$ $\frac{\alpha}{2}$ $\frac{\alpha}{2}$ O $\chi^2_{1-\alpha/2}$ $\chi^2_{\alpha/2}$
	$\sigma^2 \leqslant \sigma_0^2$			$\sigma^2 > \sigma_0^2$	$\chi^2 \geqslant \chi^2_{\alpha}(n-1)$ α O $\chi^2_{1-\alpha}$
	$\sigma^2 \geqslant \sigma_0^2$			$\sigma^2 < \sigma_0^2$	$\chi^2 \leqslant \chi^2_{1-\alpha}(n-1)$ α O $\chi^2_{1-\alpha}$

续表

条件	假设 H_0	检验统计量	统计量分布	备择假设	拒绝域
μ 已知	$\sigma^2=\sigma_0^2$	$\chi^2=\dfrac{\sum\limits_{i=1}^{n}(X_i-\mu)^2}{\sigma_0^2}$	$\chi^2(n)$	$\sigma^2\neq\sigma_0^2$	$\chi^2\geqslant\chi^2_{\alpha/2}(n)$ 或 $\chi^2\leqslant\chi^2_{1-\alpha/2}(n)$ $\frac{\alpha}{2}$ O $\chi^2_{1-\alpha/2}$ $\chi^2_{\alpha/2}$
	$\sigma^2\leqslant\sigma_0^2$			$\sigma^2>\sigma_0^2$	$\chi^2\geqslant\chi^2_{\alpha}(n)$ α O χ^2_{α}
	$\sigma^2\geqslant\sigma_0^2$			$\sigma^2<\sigma_0^2$	$\chi^2\leqslant\chi^2_{1-\alpha}(n)$ α O $\chi^2_{1-\alpha}$

2. 两个正态总体方差的假设检验

设$X_1,X_2,\cdots,X_{n_1}$是总体$X\sim N(\mu_1,\sigma_1^2)$的一个样本，$Y_1,Y_2,\cdots,Y_{n_2}$是总体$Y\sim N(\mu_2,\sigma_2^2)$的一个样本，且两样本相互独立，待检假设$H_0:\sigma_1^2=\sigma_2^2$.

(1) μ_1,μ_2 未知时，用检验统计量

$$F=S_1^2/S_2^2\sim F(n_1-1,n_2-1),$$

拒绝域为

$$F\geqslant F_{\alpha/2}(n_1-1,n_2-1)\quad 或\quad F\leqslant F_{1-\alpha/2}(n_1-1,n_2-1).$$

(2) μ_1,μ_2 已知时，用检验统计量

$$F=n_1\sum_{i=1}^{n_1}(X_i-\mu_1)^2\Big/\Big[n_2\sum_{i=1}^{n_2}(Y_i-\mu_2)^2\Big]\sim F(n_1,n_2),$$

拒绝域为

$$F\geqslant F_{\alpha/2}(n_1,n_2)\quad 或\quad F\leqslant F_{1-\alpha/2}(n_1,n_2).$$

单侧检验如表 8.11 所示.

表 8.11

条件	假设 H_0	检验统计量	统计量分布	备择假设 H_1	拒 绝 域
F 检验 μ_1,μ_2 已知	$\sigma_1^2\leqslant\sigma_2^2$	$F=\dfrac{n_1\sum_{i=1}^{n_1}(X_i-\mu_1)^2}{n_2\sum_{j=1}^{n_2}(Y_j-\mu_2)^2}$	$F(n_1,n_2)$	$\sigma_1^2>\sigma_2^2$	$F\geqslant F_{1-\alpha}(n_1,n_2)$
	$\sigma_1^2\geqslant\sigma_2^2$			$\sigma_1^2<\sigma_2^2$	$F\leqslant F_{\alpha}(n_1,n_2)$
	$\sigma_1^2=\sigma_2^2$			$\sigma_1^2\neq\sigma_2^2$	$F>F_{\alpha/2}(n_1,n_2)$ 或 $F\geqslant F_{1-\alpha/2}(n_1,n_2)$ $\frac{\alpha}{2}$ O $F_{1-\alpha/2}$ $F_{\alpha/2}$
F 检验 μ_1,μ_2 未知	$\sigma_1^2\leqslant\sigma_2^2$	$F=\dfrac{S_1^2}{S_2^2}$	$F(n_1-1,$ $n_2-1)$	$\sigma_1^2>\sigma_2^2$	$F\geqslant F_{1-\alpha}(n_1-1,n_2-1)$
	$\sigma_1^2\geqslant\sigma_2^2$			$\sigma_1^2<\sigma_2^2$	$F\leqslant F_{\alpha}(n_1-1,n_2-1)$
	$\sigma_1^2=\sigma_2^2$			$\sigma_1^2\neq\sigma_2^2$	$F>F_{1-\alpha/2}(n_1-1,n_2-1)$ 或$F\geqslant F_{\alpha/2}(n_1-1,n_2-1)$ $\frac{\alpha}{2}$ O $F_{1-\alpha/2}$ $F_{\alpha/2}$

疑难解析

1. 对于两个正态总体期望的检验,当方差未知且不相同时,若按方差未知但相等进行检验,其结果会怎样?

答 当 $\sigma_1^2=\sigma_2^2$ 时,检验统计量

$$T=\frac{\overline{X}-\overline{Y}}{S_W\sqrt{1/n_1+1/n_2}},$$

$$S_W^2=\frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}.$$

但当 $\sigma_1^2=k\sigma_2^2\ (k\neq1)$,$n_1=rn_2$ 时($r\neq1$),有

$$\frac{E(\overline{X}-\overline{Y})}{E[S_W^2(1/n_1+1/n_2)]}=1+\frac{[n_2(r+1)-1](k-1)(1-r)}{[n_2(r+k)-(k+1)](r+1)}.$$

显然仅当 $k=1$ 或 $r=1$ 时,上式等于1.

所以,当 $\sigma_1^2\neq\sigma_2^2$ 时,$|T|$ 的值会偏大或偏小,从而出现错误判断. 这时,若使样本容量 n_1 与 n_2 相等,则可使出现错误的可能性变小.

方法、技巧与典型例题分析

例 1 某维尼纶厂生产的维尼纶纤度服从正态分布,$\sigma^2\leqslant 0.048^2$. 当日随机抽取5根纤维,测得纤度如下:

$$1.55,\ 1.36,\ 1.41,\ 1.40,\ 1.32,$$

问:该日厂里生产的维尼纶纤度的方差是否正常($\alpha=0.01$)?

解 是 μ 未知、单个总体 σ^2 的假设检验,待检假设 $H_0:\sigma^2\leqslant 0.048^2$,$H_1:\sigma>0.048^2$,是 $\alpha=0.01$ 下的单侧检验.

因为 $\sigma_0^2=0.048^2$, $n=5$, $S^2=0.00788$,

所以用检验统计量 $\chi^2=\dfrac{(n-1)S^2}{\sigma_0^2}$,得

得 $$\chi^2=\frac{4\times 0.00788}{0.048^2}=13.68.$$

查表知$\chi^2_{0.01}(4)=13.3$，经比较知$\chi^2=13.68>\chi^2_{0.01}(4)=13.3$，故拒绝$H_0$，认为当日生产的维尼纶纤度方差大于$0.048^2$.

例 2 某纺织厂生产的一种细纱支数的均方差为1.2.现从当日生产的一批产品中，随机抽取了16缕进行支数测量，求得样本均方差为2.1.问：在正态总体的假定下，纱的均匀度是否变劣($\alpha=0.05$)?

解 是μ未知、单总体方差σ^2的假设检验，待检假设$H_0:\sigma^2\leqslant 1.2^2$，$H_1:\sigma^2>1.2^2$，是$\alpha=0.05$下的单侧检验.

因为 $$n=16,\quad S=2.1,\quad \sigma_0=1.2,$$

所以用检验统计量$\chi^2=\dfrac{(n-1)S^2}{\sigma_0^2}$，得

$$\chi^2=\frac{15\times 2.1^2}{1.2^2}=45.938.$$

查表知$\chi^2_{0.05}(15)=24.996$，经比较知$\chi^2=45.938>\chi^2_{0.05}(15)=24.996$，故拒绝$H_0$，认为纱的均匀度明显变劣.

例 3 自总体$X\sim N(\mu,\sigma^2)$取一容量为100的样本，测得$\bar{x}=2.7$，$\sum\limits_{i=1}^{n}(x_i-\bar{x})^2=225$.因为$\mu,\sigma^2$均未知，在$\alpha=0.05$下，检验下列假设：

(1) $H_0:\mu=3$；　　(2) $H_0:\sigma^2=2.5$.

解 (1) 是σ^2未知、单总体均值μ的假设检验.待检假设$H_0:\mu=3$，是$\alpha=0.05$下的双侧检验.

因为 $$n=100,\quad S^2=\frac{225}{99}=2.2\dot{2}\dot{7},\quad \bar{x}=2.7,$$

所以用统计量$T=\dfrac{|\overline{X}-\overline{Y}|}{S/\sqrt{n}}$，得

$$|t|=\frac{|2.7-3|}{1.508}\times\sqrt{100}=1.989.$$

查表知$t_{0.025}(99)\approx Z_{0.025}=1.96$，经比较知$|t|=1.9894>t_{0.025}(99)$

$=1.96$，拒绝 H_0. 在实际问题中，如果两数过于接近，则不宜作出结论，应提出重新抽样，再作一次检验的要求.

(2) 由于 $\sum_{i=1}^{n}(x_i-\bar{x})$ 已知，可用检验统计量 $\chi^2=\frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{\sigma^2}$，待检假设 $H_0:\sigma^2=2.5$，是双侧检验.

$$\chi^2=\frac{225}{2.5}=90.$$

查表知 $\quad \chi^2_{0.025}(99)=129.56,\quad \chi^2_{0.975}(99)=74.22,$

而 $$\chi^2_{0.975}(99)<\chi^2=90<\chi^2_{0.025}(99).$$

故接受 H_0，认为方差 $\sigma^2=2.5$.

例4 包装机包装食盐，设每袋盐的净重服从正态分布，规定每袋标准重量为 500 g，标准差不超过 10 g. 某日开工后，随机抽取 9 袋，测得净重(单位:g)如下：

497, 507, 510, 475, 515, 484, 488, 524,491,

在 $\alpha=0.05$ 下检验假设：

(1) $H_0:\mu=500$； (2) $H_0:\sigma\leqslant 10, H_1:\sigma>10$.

解 (1) 是 σ^2 未知、单总体均值 μ 的假设检验. 待检假设 $H_0:\mu=500$，是 $\alpha=0.05$ 下的双侧检验.

因为 $\quad n=9,\quad \bar{x}=499,\quad S=16.03,$

所以用检验统计量 $T=\frac{|\overline{X}-\overline{Y}|}{S/\sqrt{n}}$，得

$$|t|=\frac{|499-500|}{16.03}\times\sqrt{9}=0.187.$$

查表知 $t_{0.025}(8)=2.306$，经比较知 $|t|=0.187<t_{0.025}(8)=2.306$，故接受 H_0，认为该天生产的食盐每袋净重是 500 g.

(2) 是 μ 未知、单总体方差的假设检验，待检假设 $H_0:\sigma\leqslant 10$，$H_1:\sigma>10$，是 $\alpha=0.05$ 下的单侧检验.

因为 $\quad n=9,\quad S=16.301,\quad \sigma_0=10,$

所以用检验统计量 $\chi^2=\dfrac{(n-1)S^2}{\sigma_0^2}$，得

$$\chi^2=\frac{8\times 16.301^2}{10^2}=21.258.$$

查表知 $\chi^2_{0.05}(8)=15.507$，经比较知 $\chi^2=21.258>\chi^2_{0.05}(8)=15.507$，故拒绝 H_0，认为该天生产的食盐每袋净重的标准差大于 10 g.

例5 设随机变量 X 与 Y 相互独立，$X\sim N(\mu_1,\sigma_1^2)$，$Y\sim N(\mu_2,\sigma_2^2)$. $X_1,X_2,\cdots,X_{16}$是X 的一个样本，$Y_1,Y_2,\cdots,Y_{10}$是Y 的一个样本，测得数据

$$\sum_{i=1}^{16}x_i=84,\quad \sum_{i=1}^{16}x_i^2=563,\quad \sum_{i=1}^{10}y_i=18,\quad \sum_{i=1}^{10}y_i^2=72.$$

(1) 分别求 μ_1,μ_2 的矩估计值；

(2) 分别求 σ_1^2,σ_2^2 的极大似然估计值；

(3) 在显著性水平 $\alpha=0.05$ 下检验假设

$$H_0:\sigma_1^2\leqslant\sigma_2^2,\quad H_1:\sigma_1^2>\sigma_2^2.$$

解 (1) 用样本一阶原点矩估计总体一阶矩，即得 μ_1 和 μ_2 的矩估计值为

$$\hat{\mu}_1=\bar{x}=\frac{1}{16}\sum_{i=1}^{16}x_i=5.25,\quad \hat{\mu}_2=\bar{y}=\frac{1}{10}\sum_{i=1}^{10}y_i=1.8.$$

(2) 正态总体 $X\sim N(\mu,\sigma^2)$的参数 σ^2 的极大似然估计量为 $\hat{\sigma}^2=\dfrac{1}{n}\sum\limits_{i=1}^{n}(X_i-\overline{X})^2$，因此 σ_1^2 和 σ_2^2 的极大似然估计值为

$$\hat{\sigma}^2=\frac{1}{16}\sum_{i=1}^{16}(x_i-\bar{x})^2=\frac{1}{16}\left(\sum_{i=1}^{n}x_i^2-16\bar{x}^2\right)=7.63,$$

$$\hat{\sigma}_2^2=\frac{1}{10}\sum_{i=1}^{10}(y_i-\bar{y})^2=\frac{1}{10}\left(\sum_{i=1}^{n}y_i^2-10\bar{y}^2\right)=3.96.$$

(3) 是μ_1,μ_2 未知、双总体方差的假设检验，待检假设$H_0:\sigma_1^2\leqslant\sigma_2^2$；$H_1:\sigma_1^2>\sigma_2^2$，是在$\alpha=0.05$ 下的单侧检验.

因为 $$S_1^2=\frac{1}{15}\sum_{i=1}^{16}(x_i-\bar{x})^2=8.13,$$

$$S_2^2=\frac{1}{9}\sum_{i=1}^{10}(y_i-\bar{y})^2=4.4.$$

所以用检验统计量 $F=S_1^2/S_2^2$，得

$$F=\frac{8.13}{4.4}=1.85.$$

查表知 $F_{0.05}(15,9)=3.01$，经比较知 $F=1.85<F_{0.05}(15,9)=3.01$，故接受 H_0，认为 σ_1^2 不比 σ_2^2 大.

例6 对某种金属的熔点作了四次测定，数据(单位：℃)如下：

$$1269,\ 1271,\ 1263,\ 1265,$$

假定数据服从正态分布. 在显著性水平 $\alpha=0.05$ 下，检验假设测定值的均方差不大于2是否成立.

解 设熔点 $X\sim N(\mu,\sigma^2)$，是 μ 未知、单总体方差的假设检验，待检假设 $H_0:\sigma^2\leqslant 4, H_1:\sigma^2>4$，是 $\alpha=0.05$ 下的单侧检验.

因为 $n=4, \bar{x}=12.67, S^2=13.3$，所以用检验统计量 $\chi^2=\frac{(n-1)S^2}{\sigma_0^2}$，得

$$\chi^2=\frac{3\times 13.3}{4}=10.$$

查表知 $\chi^2_{0.05}(3)=9.378$，经比较知 $\chi^2=10>\chi^2_{0.05}(3)=9.378$，故拒绝 H_0，认为测定值的均方差大于2 ℃.

例7 两台机床加工同一种零件，分别取6个和9个零件测量其长度(单位：mm)，计算得

$$S_1^2=0.345,\quad S_2^2=0.357.$$

假定零件长度服从正态分布，是否可认为两台机床加工的零件尺寸的方差无显著差异($\alpha=0.05$)？

解 是 μ_1,μ_2 均未知、双总体的方差的假设检验. 待检假设 $H_0:\sigma_1^2=\sigma_2^2$，是 $\alpha=0.05$ 下的双侧检验.

因为 $S_1^2=0.345, S_2^2=0.357$，所以用检验统计量 $F=S_1^2/S_2^2$，得

$$F=\frac{0.345}{0.357}=0.966.$$

查表知 $F_{0.025}(5,8)=4.82, F_{0.975}(5,8)=1/F_{0.025}(8,5)=0.148$，经比较知 $F_{0.975}(5,8)=4.82<F=0.966<F_{0.025}(5,8)=0.148$，故接受 H_0，认为两机床加工的零件尺寸无显著差异.

例 8 某实验室有 A、B 两种仪器，测量某一物体长度（单位：mm）7 次和 10 次，得数据如表 8.12 所示. 在 $\alpha=0.05$ 下，能否认为仪器B 的精度比仪器A 的精度高？

表 8.12

A	97	102	103	96	100	101	100			
B	100	101	103	98	97	99	102	101	98	101

解 物体的长度一般服从正态分布. 但 μ_1, μ_2 均未知，可认为方差小的精度高. 待检假设 $H_0: \sigma_1^2>\sigma_2^2, H_1: \sigma_1^2 \leqslant \sigma_2^2$，是单侧检验.

计算得 $S_1^2=6.4762, S_2^2=3.7778$，用检验统计量 $F=S_1^2/S_2^2$，得

$$F=\frac{6.4762}{3.7778}=1.714.$$

查表知 $F_{0.05}(6,9)=3.37$，经比较知 $F=1.714<F_{0.05}(6,9)=3.37$，故拒绝 H_0，认为仪器B 的精度比仪器A 高.

例 9 对两批同型号的电子元件各抽取 6 件进行测试，得电阻（单位：Ω）数据如表 8.13 所示. 设元件的电阻总体分别服从 $N(\mu_1, \sigma_1^2)$ 和 $N(\mu_2, \sigma_2^2)$，且相互独立. 试在 $\alpha=0.05$ 下检验两批元件的电阻有无显著差异.

表 8.13

A 批	0.140	0.138	0.143	0.141	0.144	0.137
B 批	0.135	0.140	0.142	0.136	0.138	0.141

解 待检假设 $H_0: \mu_1=\mu_2$. 但由于 σ_1^2, σ_2^2 未知，考虑方差齐性，必须先检验 σ_1^2 是否等于 σ_2^2，故提出假设 $H'_0: \sigma_1^2=\sigma_2^2$，是双侧检验.

因为

$\bar{x}=0.1405$， $\bar{y}=0.1387$， $S_1=0.0027$， $S_2=0.0026$，
所以用检验统计量 $F=S_1^2/S_2^2$，得

$$F=\frac{0.0027^2}{0.0026^2}=1.078.$$

查表知 $F_{0.025}(5,5)=7.15$， $F_{0.975}(5,5)=\frac{1}{F_{0.025}(5,5)}=0.14$，经比较知 $0.14<F=1.078<7.15$，故接受 H_0'，认为两批电子元件总体的方差相等.

由 $\sigma_1^2=\sigma_2^2=\sigma^2$ 但 σ^2 未知的条件，检验两总体均值是否相等，可采用检验统计量 T；待检假设 $H_0:\mu_1=\mu_2$，是双侧检验.

用检验统计量

$$T=\frac{|\bar{X}-\bar{Y}|}{\sqrt{\frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}}}\cdot\frac{1}{\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}$$

得

$$|t|=\frac{0.1405-0.1387}{\sqrt{\frac{(6-1)\times0.0027^2+(6-1)\times0.0026^2}{6+6-2}}}\cdot\frac{1}{\sqrt{\frac{1}{6}+\frac{1}{6}}}$$

$=1.1763.$

查表知 $t_{0.025}(12-2)=2.2281$，经比较知 $|t|=1.1763<t_{0.025}(10)=2.2281$，故接受 H_0，认为两批元件的电阻无显著差异.

例 10 某化工厂为了提高一种化工产品的得率（质量分数，%），采用了两种方案，进行各10 次试验，测得数据如表8.14 所示. 假设得率服从正态分布，试在 $\alpha=0.05$ 下检验甲方案得率是否高于乙方案.

表 8.14

甲	68.1	62.4	64.3	64.7	68.4	66.0	65.5	66.7	67.3	66.2
乙	69.1	71.0	69.1	70.0	69.1	69.1	67.3	70.2	72.1	67.3

解 $\mu_1,\mu_2,\sigma_1^2,\sigma_2^2$ 均未知，故应先作方差齐性检验. 待检假设 $H_0':\sigma_1^2=\sigma_2^2$，是 $\alpha=0.50$ 下的双侧检验（因为此时不保护原假设 H_0'，

故 α 取0.5).

因为 $S_1^2=3.3511$，$S_2^2=2.2244$，$n_1=n_2=10$，
所以用检验统计量 $F=S_1^2/S_2^2$,得

$$F=\frac{3.3511}{2.2244}=1.51.$$

查表知 $F_{0.25}(9,9)=0.629$,$F_{0.75}(9,9)=1.59$,经比较知 $F_{0.25}(9,9)=0.629<F=1.51<F_{0.75}(9,9)=1.59$,故接受 H_0',认为两种方案得率的方差相等.

在 $\sigma_1^2=\sigma_2^2=\sigma^2$ 但 σ^2 未知的条件下,进行两总体均值的假设检验,待检假设 $H_0:\mu_1\geqslant\mu_2$,是 $\alpha=0.05$ 下的单侧检验.

因为 $\bar{x}=65.96$，$\bar{y}=69.43$，$n_1=n_2=10$，
所以用检验统计量 T,得

$$t=\frac{65.96-69.43}{\sqrt{\frac{(10-1)\times 3.3511+(10-1)\times 2.2244}{10+10-2}}}\cdot\frac{1}{\sqrt{\frac{1}{10}+\frac{1}{10}}}$$
$$=-4.6472.$$

查表知 $t_{0.95}(18)=-1.7341$,经比较知 $t=-4.6472<t_{0.95}(18)=-1.7341$,故拒绝 H_0,认为乙方案比甲方案可明显提高得率.

例11 某农业试验站为了研究一种肥料对提高农作物产量的效果,分别取6块和7块条件相同的小区作对比试验,得数据如表8.15所示.设农作物产量服从正态分布,试在 $\alpha=0.10$ 下检验这种化肥对提高农作物产量的效果是否显著.

表8.15

施肥	34	35	32	33	34	30	
不施肥	29	27	32	31	28	32	31

解 是 $\mu_1,\mu_2,\sigma_1^2,\sigma_2^2$ 均未知的条件下的假设检验,需先检验 $H_0':\sigma_1^2=\sigma_2^2$,是双侧检验.

因为 $n_1=6$，$n_2=7$，$S_1^2=3.2$，$S_2^2=4$，

所以用统计量$F=S_1^2/S_2^2$,得

$$F=\frac{3.2}{4}=0.8.$$

查表知$F_{0.05}(5,6)=4.39$,$F_{0.95}(5,6)=\frac{1}{F_{0.05}(6,5)}=0.202$,经比较知$F_{0.95}(5,6)=0.202<F=0.8<F_{0.05}(5,6)=4.39$,故接受$H_0$,认为两总体的方差相等.

检验假设$H_0:\mu_1\leqslant\mu_2$,是方差未知但相等条件下的单侧检验.

因为$\bar{x}=33$,$\bar{y}=30$,所以用检验统计量T,得

$$t=\frac{33-30}{\sqrt{\frac{(6-1)\times 32+(7-1)\times 4}{6+7-2}}}\cdot\frac{1}{\sqrt{\frac{1}{6}+\frac{1}{7}}}=2.828.$$

查表知$t_{0.10}(11)=1.363$,经比较知$t=2.828>t_{0.10}(11)=1.363$,故拒绝$H_0$,认为化肥对提高农作物产量的效果是明显的.

例12 从某学院学生的经常参加锻炼和不经常参加锻炼的男生中各随机抽取50名,测得平均身高(单位:cm)$\bar{x}=174.34$,$\bar{y}=172.42$.设身高服从正态分布,且已知$\sigma_1=5.35$,$\sigma_2=6.11$.问:经常参加锻炼的学生是否高于不锻炼的学生($\alpha=0.05$)?

解 σ_1^2,σ_2^2已知,所以不必作方差齐性检验,待检假设$H_0:\mu_1\leqslant\mu_2$;$H_1:\mu_1>\mu_2$,是单侧检验.

因为

$$\bar{x}=174.34,\quad \bar{y}=172.42,\quad n_1=n_2=50,\quad \sigma_1=5.35,\quad \sigma_2=6.11,$$

所以用统计量$U=\frac{\overline{X}-\overline{Y}}{\sqrt{\sigma_1^2/n_1+\sigma_2^2/n_2}}$,得

$$u=\frac{174.34-172.42}{\sqrt{\frac{5.35^2}{50}+\frac{6.11^2}{50}}}=1.67.$$

查表知$Z_{0.95}=1.64$,经比较知$u=1.67>Z_{0.95}=1.64$,故拒绝H_0,认为经常锻炼的学生高于不经常锻炼的学生.

此例再次说明,方差是否相等很重要.

第三节　总体分布的假设检验

主 要 内 容

总体分布的假设检验是针对分布本身，而不是针对分布中参数的检验，又称为非参数检验.

1. χ^2 拟合优度检验法

(1) 总体 X 只取有限个值的情形　设总体 X 是仅取有限个 (k) 值的离散型随机变量，样本为 $X_1, X_2, \cdots, X_n$，检验假设 H_0：总体 X 的分布律为 $P\{X=i\}=P_i,\ i=1,2,\cdots,k$.

由皮尔逊(K. Pearson)定理可知：当 H_0 成立时，如 $n\to\infty$ 时，近似地有检验量 $\chi^2 \sim \chi^2(k-1)$.

因此，对于给定的显著性水平 α，如果

$$\chi^2=\sum_{i=1}^{k}\frac{(f_i-np_i)^2}{np_i}\geqslant\chi^2_\alpha(k-1),$$

则拒绝 H_0，否则接受 H_0.

要求 $n\geqslant 50$，$np_i\geqslant 5$. 若 $np_i<5$，应适当合并事件，使 $np_i\geqslant 5$.

这里，频率 f_i/n 是在 n 次试验中，事件 A_i 出现的频率.

(2) 总体 X 取无限多个值的情形　设总体的分布函数为 $F(x)$，$X_1, X_2, \cdots, X_n$ 为来自 X 的一个样本，检验假设 H_0：$F(x)=F_0(x)$（或 $f(x)=f_0(x)$）.

由皮尔逊定理可知：当 H_0 为真时，若 n 充分大($n\geqslant 50$)，则不论总体 X 属于什么分布，$\chi^2=\sum\limits_{i=1}^{n}\frac{(f_i-np_i)^2}{np_i}$ 总是近似服从 $\chi^2(k-r-1)$ 分布，其中 r 是被估计参数的个数.

因此，在显著性水平 α 下，当

$$\chi^2=\sum_{i=1}^{n}\frac{(f_i-np_i)^2}{np_i}>\chi_\alpha^2(k-r-1)$$

时，拒绝 H_0，否则接受 H_0.

当 $np_i<5$ 时，应适当合并区间，使 $np_i\geqslant 5$.

2. 秩和检验法

秩和检验法是一种用样本秩代替样本值的检验方法，用来检验两总体的分布函数是否相同.

设两总体 X 和 Y 的分布函数分别为 $F(x)$ 和 $F(y)$，$X_1,X_2,\cdots,X_{n_1}$ 和 $Y_1,Y_2,\cdots,Y_{n_2}$ 分别是 X 和 Y 的两个相互独立的样本. 检验假设 $H_0:F(x)=F(y)$.

(1) 将样本观察值按由小到大的次序编号，规定数据的序数 k 为该数的秩.

(2) 将分别属于第一总体和第二总体的样本观察值的秩相加，得到第一样本和第二样本的秩和 R_1 和 R_2.

(3) 取容量小的样本的秩和为检验统计量 T.

(4) 由样本容量 n_1 和 n_2 及检验水平 α，通过秩和检验表($\alpha=0.05$)，查出 C_1 和 C_2.

若 $C_1<T<C_2$，则接受 H_0；若 $T<C_1$ 或 $T>C_2$，则拒绝 H_0.

注意，括号内数字为样本容量(n_1,n_2)，括号下的两对数字为临界点(对应不同的犯第一类错误的概率)，右边小数为犯第一类错误的概率.

当 n_1,n_2 大于 10 时，用检验统计量 U，此时

$$T\dot{\sim}N\left(\frac{n_1(n_1+n_2+1)}{2},\frac{n_1n_2(n_1+n_2+1)}{12}\right),$$

统计量为

$$U=\frac{T-n_1(n_1+n_2+1)/2}{\sqrt{n_1n_2(n_1+n_2+1)/12}}\sim N(0,1).$$

疑难解析

1. χ^2 拟合优度检验法的基本思想是什么？

答 χ^2 拟合优度检验法的基本思想是：将随机试验的可能结果的全体Ω，分为k个互不相容的事件$A_1,A_2,\cdots,A_k$，使$\sum_{i=1}^{k}A_k=\Omega$，$A_iA_j=\varnothing, i\neq j\ (i,j=1,2,\cdots,k)$. 在假设$H_0$下，计算概率$p_i=P(A_i)$（或者用极大似然估计法估计$\hat{p}_i$）. 在$n$次试验中，统计事件$A_i$出现的频率，比较频率$f_i/n$与概率$p_i$（或$\hat{p}_i$）. 在$H_0$为真，且试验次数很大时，频率与概率的差异应很小. 由此提出检验统计量

$$\chi^2=\sum_{i=1}^{k}\frac{(f_i-np_i)^2}{np_i}\left(\text{或}\ \chi^2=\sum_{i=1}^{k}\frac{(f_i-n\hat{p}_i)^2}{n\hat{p}_i}\right)$$

检验假设H_0.

一般要求$n\geqslant 50$.

2. 怎样确定χ^2 拟合优度检验中的$F_0(x)$？还应注意哪些问题？

答 $F_0(x)$可由总体的具体情形初步确定，然后用极大似然估计法估计参数的值，用估计值代替参数值. 具体的做法是：将区间$(-\infty,+\infty)$分成k个互不相交的区间$(-\infty,a_1]$，$(a_1,a_2]$，$\cdots$，$(a_{k-1},+\infty]$（一般可取$k=5\sim10$）；再求出落在第i个小区间内的样本值个数$f_i\ \ (i=1,2,\cdots,k)$，得到频率f_i/n，记$a_0=-\infty$，$a_k=+\infty$；然后令$p_i=F(a_i)-F(a_{i-1})$，$i=1,\cdots,k$，用χ^2统计量检验.

还应注意的是：(1)若总体分布含有r个未知参数，则临界值由$\chi^2(k-r-1)$确定；(2)将$(-\infty,+\infty)$分为k个区间的分法是任意的，但一般取$k=5\sim10$. 当分布是对称时，区间也最好取为对称的；(3)n应当较大($n\geqslant 0$)，则临界域的结论较可靠.

3. 秩和检验的基本思想是什么？

答 秩和检验法的基本思想是：取样本容量小的秩和作为统计量T，根据统计量的值来求检验问题的拒绝域. 威尔・柯克逊认为，

两个样本秩和总和$R_1+R_2=\frac{1}{2}(n_1+n_2)(n_1+n_2+1)$. 当$H_0$成立时,可以认为两独立样本实际上来自同一总体,因此X_i和Y_i的值出现在排列的每一位置上的可能性相同. 故样本容量小的样本的各元素应分散在排列中,集中在排列前边或后面的可能性很小,所以秩和T的值不应太大或太小. 于是当T较大或较小时,拒绝H_0.

方法、技巧与典型例题分析

一、χ^2拟合优度检验法

χ^2拟合优度检验法检验总体分布,要初步提出一个关于总体的假设,这一假设大多利用极大似然估计法作出,但也可以借助图形(如直方图上端连线)确定,方法较多. 读者应多看一些参考书籍,提高自己的能力.

例1 一枚骰子掷了100次,得结果如表8.16所示,在$\alpha=0.05$下,检验这枚骰子是否均匀.

表8.16

点　数	1	2	3	4	5	6
频数f_i	13	14	20	17	15	21

解 用X表示骰子掷出的点数,$P\{X=i\}=p_i$, $i=1,2,\cdots,6$. 如果骰子是均匀的,则$p_i=1/6$, $i=1,2,\cdots,6$. 待检假设

$$H_0: p_i=1/6.$$

计算检验统计量$\chi^2=\sum_{i=1}^{n}\frac{(f_i-np_i)^2}{np_i}$的值,得

$$\chi^2=\left[\left(13-\frac{100}{6}\right)^2+\left(14-\frac{100}{6}\right)^2+\left(20-\frac{100}{6}\right)^2+\left(17-\frac{100}{6}\right)^2+\left(15-\frac{100}{6}\right)^2+\left(21-\frac{100}{6}\right)^2\right]\div\frac{100}{6}$$
$$=3.2.$$

查表知 $\chi^2_{0.05}(6-1)=11.071$，经比较知 $\chi^2=3.2<\chi^2_{0.05}(5)=11.071$，故接受 H_0，认为骰子是均匀的.

例 2 某厂近年来发生了 63 次事故，按星期几统计如表 8.17 所示，问：事故的发生是否与星期几有关($\alpha=0.05$)？

表 8.17

星 期	一	二	三	四	五	六
频数 f_i	9	10	11	8	13	12

解 设 X 为事故发生在星期 i，$i=1,2,\cdots,6$，$P\{X=i\}=p_i$. 若事故的发生与星期几无关，则 $p_i=1/6$，$i=1,2,\cdots,6$. 待检假设

$$H_0: p_i=1/6.$$

计算检验统计量 $\chi^2=\sum\limits_{i=1}^{n}\dfrac{(f_i-np_i)^2}{np_i}$ 的值，得

$$\chi^2=\left[\left(9-\frac{63}{6}\right)^2+\left(10-\frac{63}{6}\right)^2+\left(11-\frac{63}{6}\right)^2+\left(8-\frac{63}{6}\right)^2+\left(13-\frac{63}{6}\right)^2+\left(12-\frac{63}{6}\right)^2\right]\div\frac{63}{6}$$

$$=1.67.$$

查表知 $\chi^2_{0.05}(5)=11.07$，经比较知 $\chi^2=1.67<\chi^2_{0.05}(5)=11.07$，故接受 H_0，认为事故的发生与星期几没有关系.

例 3 孟德尔(Mendel)在豌豆试验中对 10 棵豌豆株统计了黄色豌豆与青色豌豆颗数，得：黄色豌豆 355，青色豌豆 123. 孟德尔的理论认为，青、黄豌豆比为 1∶3. 试在 $\alpha=0.05$ 下检验这一假设.

解 设 $P\{黄\}=3/4$，$P\{青\}=1/4$. 待检假设 H_0：$P(黄)=3/4$，$P\{青\}=1/4$.

计算检验统计量 $\chi^2=\sum\limits_{i=1}^{n}\dfrac{(f_i-np_i)^2}{np_i}$ 的值，得

$$\chi^2=\frac{(355-478\times 3/4)^2}{478\times 3/4}+\frac{(123-478\times 1/4)^2}{478\times 1/4}=0.1367.$$

查表知 $\chi^2_{0.05}(1)=3.841$，经比较知 $\chi^2=0.1367<\chi^2_{0.05}(1)=3.841$，故接受 H_0，认为青、黄豌豆比为 1∶3.

例 4 某船厂的历史资料显示，生产的农用船销往 A、B、C、D、E 地区的比例为 20%，28%，8%，12%，32%. 在今年生产的农用船中观察了 500 艘，发现销往上述地区的分别为 120，123，43，66，148 艘，试在 $\alpha=0.05$ 下检验销售比例是否改变.

解 设 X 为销往五个地区事件，则农用船销往不同地区的概率如表 8.18 所示，待检假设 $H_0:P\{X=i\}=p_i$.

表 8.18

X	A	B	C	D	E
p_i	0.20	0.28	0.08	0.12	0.32

计算检验统计量 $\chi^2=\sum_{i=1}^{n}\frac{(f_i-np_i)^2}{np_i}$ 的值，得

$$\chi^2=\frac{(120-500\times0.2)^2}{500\times0.2}+\frac{(123-500\times0.28)^2}{500\times0.28}+\frac{(43-500\times0.08)^2}{500\times0.08}+\frac{(66-500\times0.12)^2}{500\times0.12}+\frac{(148-500\times0.32)^2}{500\times0.32}=7.789.$$

查表知 $\chi^2_{0.05}(4)=9.488$，经比较知 $\chi^2=7.789<\chi^2_{0.05}(4)=9.488$，故接受 H_0，认为销售比例与历年无显著变化.

例 5 从总体 X 中抽取一个容量为 80 的样本，得频数分布如下表 8.19 所示，试在 $\alpha=0.025$ 下检验 H_0:X 的概率密度

$$f(x)=\begin{cases}2x, & 0<x<1,\\ 0. & \text{其它}.\end{cases}$$

表 8.19

区间	$\left(0,\frac{1}{4}\right]$	$\left(\frac{1}{4},\frac{1}{2}\right]$	$\left(\frac{1}{2},\frac{3}{4}\right]$	$\left(\frac{3}{4},1\right]$
频数	6	18	20	36

解 因为

$$p_i=P\left\{\frac{i-1}{4}<X\leqslant\frac{i}{4}\right\}=\int_{(i-1)/4}^{i/4}2x\mathrm{d}x=\frac{i^2-(i-1)^2}{16},$$

$$i=1,2,3,4,$$

待检假设 $H_0:X\sim f(x)=\begin{cases}2x, & 0<x<1,\\ 0, & \text{其它}.\end{cases}$

列计算表如表8.20所示.算得

$$\chi^2=\sum_{i=1}^{4}\frac{(f_i-np_i)^2}{np_i}=1.83.$$

表8.20

i	区间	f_i	p_i	np_i	f_i-np_i	$(f_i-np_i)^2/np_i$
1	$\left(0,\frac{1}{4}\right]$	6	0.0625	5	1	0.20
2	$\left(\frac{1}{4},\frac{1}{2}\right]$	18	0.1875	15	3	0.60
3	$\left(\frac{1}{2},\frac{3}{4}\right]$	20	0.3125	25	−5	1.00
4	$\left(\frac{3}{4},1\right]$	36	0.4375	35	1	0.03

查表知 $\chi^2_{0.025}(3)=9.348$,经比较知

$$\chi^2=1.83<\chi^2_{0.025}(3)=9.348,$$

故接受 H_0,认为 X 的概率密度为

$$f(x)=\begin{cases}2x, & 0<x<1,\\ 0, & \text{其它}.\end{cases}$$

例6 考察某地区110 kV电网在某天内电压的波动情况,记录了当天的100个电压数据(单位:kV),经分组整理后如表8.21所示,且得 $\bar{x}=109.52$, $S=1.88$.试问:该电网电压是否服从正态分布($\alpha=0.05$)?

表 8.21

区间	频数	区间	频数
$(-\infty,106.55]$	6	$(109.55,110.55]$	23
$(106.55,107.55]$	8	$(110.55,111.55]$	15
$(107.55,108.55]$	13	$(111.55,112.55]$	9
$(108.55,109.55]$	21	$(112.55,+\infty]$	5

解 依题意，待检假设 $H_0: X \sim N(\mu,\sigma^2)$. 参数 μ,σ^2 未知，由极大似然估计法，有 $\hat{\mu}=\overline{X}=109.52$，$\hat{\sigma}^2\approx S^2=1.88^2$，用估计值代替参数，所以待检假设 $H_0: X\sim N(109.52,1.88^2)$. 由

$$\hat{p}_i=P\{a_{i-1}<X\leqslant a_i\}=\Phi\left(\frac{a_i-\overline{X}}{\sigma_i}\right)-\Phi\left(\frac{a_{i-1}-\overline{X}}{\sigma_i}\right)$$

$$(i=1,2,\cdots,8;a_0=-\infty,a_8=+\infty),$$

将计算结果列于表 8.22 中. 算得

$$\chi^2=\sum_{i=1}^{8}\frac{(f_i-n\hat{p}_i)^2}{n\hat{p}_i}=0.937.$$

表 8.22

i	f_i	p_i	np_i	$(f_i-n\hat{p}_i)^2$	$(f_i-n\hat{p}_i)^2/n\hat{p}_i$
1	6	0.057	5.7	0.09	0.016
2	8	0.899	8.99	0.98	0.109
3	13	0.1546	15.46	6.05	0.391
4	21	0.2045	20.45	0.303	0.015
5	23	0.2028	20.28	7.398	0.365
6	15	0.1511	15.11	0.012	0.001
7	9	0.0864	8.64	0.130	0.015
8	5	0.0537	5.37	0.137	0.025

因为$k=8$，$r=2$，$k-r-1=5$，查表知$\chi^2_{0.05}(5)=11.071$，经比较知$\chi^2=0.937<\chi^2_{0.05}(5)=11.071$，故接受$H_0$，认为该日电网电压服从正态分布$N(109.52,1.88^2)$.

例7 为了考察某传呼台在中午12时至13时电话呼错的次数，统计了200 d的记录如表8.23所示，问：在显著性水平$\alpha=0.25$下，能否认为总体服从泊松分布？

表8.23

呼错次数	0	1	2	3	4
频数f_i	109	65	22	3	1

解 待检假设H_0：$X\sim\pi(\lambda)$ $(\lambda>0)$.

用极大似然法估计$\hat{\lambda}=\overline{X}$，计算估计值，得

$$\hat{\lambda}=\overline{x}=\frac{1}{200}(0\times109+1\times65+2\times22+3\times3+4\times1)=0.61.$$

由$X\sim\pi(0.61)$，得

$$\hat{p}_0=e^{-0.61}\frac{0.61^0}{0!}=0.543,\quad \hat{p}_1=0.331,$$

$$\hat{p}_2=0.101,\quad \hat{p}_3=0.021,\quad \hat{p}_4=0.004.$$

又$n\hat{p}_4=0.8<5$，与$\hat{p}_3$合并，得到如表8.24所示结果.算得

$$\chi^2=\sum_{i=1}^{n}\frac{(f_i-n\hat{p}_i)^2}{n\hat{p}_i}=0.384.$$

表8.24

x_i	f_i	$\hat{p}_i$	$n\hat{p}_i$	$(f_i-n\hat{p}_i)^2$	$(f_i-n\hat{p}_i)^2/n\hat{p}_i$
0	109	0.543	108.6	0.16	0.00147
1	65	0.331	66.2	1.44	0.02175
2	22	0.101	20.2	3.24	0.16039
3	3	0.021	5	1	0.2
4	1	0.004			

查表知$\chi^2_{0.25}(4-1-1)=\chi^2_{0.25}(2)=2.773$，经比较知$\chi^2=0.384<$

$\chi^2_{0.25}(2)=2.773$,故接受H_0,认为传呼台呼错次数服从泊松分布.

例8 每次检查产品时,都抽取10件产品来检查.统计100次的检查结果,得到每10件产品中次品数的分布如表8.25所示,试用 χ^2 拟合优度检验法检验次品数总体 X 服从二项分布($\alpha=0.05$)?

表8.25

次品数x_i	0	1	2	3	4	5	$\geqslant 6$
频数f_i	32	45	17	4	1	1	0

解 待检假设$X\sim B(n,p)$.由极大似然估计法,$\hat{p}=\overline{x}/n$.经计算,得

$$\hat{p}=\frac{1}{100\times 10}(45+2\times 17+3\times 4+4\times 1+5\times 1)=\frac{1}{10},$$

故待检假设$H_0:X\sim B(10,1/10)$.

$$P\{X=i\}=C_{10}^{i}\times(1/10)^{i}\times(9/10)^{10-i},\ i=1,2,\cdots,10.$$

为了使$n\hat{p}_i\geqslant 5$,将$x_i=3,4,5$合并,于是$k=4$,$r=1$.计算χ^2的观察值(计算数据不再列出),得

$$\chi^2=\sum_{i=0}^{3}\frac{(f_i-n\hat{p}_i)^2}{n\hat{p}_i}=1.69.$$

查表知$\chi^2_{0.05}(4-1-1)=5.99$,经比较知$\chi^2=1.69<\chi^2_{0.05}(2)=5.99$,故接受$H_0$,认为10件产品中的次品数服从二项分布$B(10,1/10)$.

例9 在$\pi=3.14159\cdots$的前800位小数中各数字出现的次数如表8.26所示,试用χ^2拟合优度检验法检验各数字的分布是否服从均匀分布($\alpha=0.05$).

表8.26

数	0	1	2	3	4	5	6	7	8	9
频数	74	92	83	79	80	73	99	75	76	91

解 以X记π的小数部分出现的数字,$P\{X=i\}=p_i$,$i=$

0,1,…,9. 若 X 服从均匀分布，则 $p_i=1/10$，故待检假设

$$H_0: p_i=1/10,\quad i=0,1,\cdots,9.$$

列出计算结果如表 8.27 所示. 算得

$$\chi^2=\sum_{i=0}^{9}\frac{(f_i-np_i)^2}{np_i}=5.125.$$

表 8.27

i	0	1	2	3	4	5	6	7	8	9
f_i	74	92	83	79	80	73	99	75	76	91
$\lvert f_i-np_i\rvert$	6	12	8	1	0	7	3	5	4	11
$\frac{(f_i-np_i)^2}{np_i}$	0.45	1.80	0.1125	0.0125	0	0.6125	0.1125	0.3125	0.20	1.5125

查表知 $\chi^2_{0.05}(10-1)=16.919$，经比较知 $\chi^2=5.125<\chi^2_{0.05}(9)=16.919$，故接受 H_0，认为 π 的小数部分各数字的分布服从均匀分布.

例 10 某运动员用手枪对 100 个靶各射击 10 发子弹，记录射击的结果如表8.28 所示，试用 χ^2 拟合优度检验法检验射击结果是否服从二项分布($\alpha=0.05$).

表 8.28

命中枪数	0	1	2	3	4	5	6	7	8	9	10
靶　数	0	2	4	10	22	26	18	12	4	2	0

解 设射击命中枪数 $X\sim B(n,p)$，分布律

$$p_n(k)=C_n^k p^k(1-p)^{n-k},\ k=0,1,\cdots,10.$$

由极大似然估计法，$\hat{p}=\overline{X}/n=5/10=0.5$. 因为 np_i 应不小于 5，将 0,1,2 和 8,9,10 合并，因而有 $N=6$. 列出计算结果如表 8.29 所示. 算得

$$\chi^2=\sum_{i=1}^{7}\frac{(f_i-np_i)^2}{np_i}=0.879.$$

表 8.29

N	x_i	f_i	p_i	np_i	f_i-np_i	$(f_i-np_i)^2/np_i$
1	0,1,2	6	$56\times\left(\frac{1}{2}\right)^{10}$	5.5	0.5	0.045
2	3	10	$120\times\left(\frac{1}{2}\right)^{10}$	11.7	−1.7	0.247
3	4	22	$210\times\left(\frac{1}{2}\right)^{10}$	20.5	1.5	0.110
4	5	26	$252\times\left(\frac{1}{2}\right)^{10}$	24.6	1.4	0.080
5	6	18	$210\times\left(\frac{1}{2}\right)^{10}$	20.5	−2.5	0.305
6	7	12	$120\times\left(\frac{1}{2}\right)^{10}$	11.7	0.3	0.008
7	8,9,10	6	$56\times\left(\frac{1}{2}\right)^{10}$	5.5	0.5	0.084

查表知$\chi^2_{0.05}(7-1-1)=11.071$，经比较知$\chi^2=0.879<\chi^2_{0.05}(5)=11.071$，故接受$H_0$，认为该运动员射击命中枪数服从二项分布.

例 11 在一批灯泡中做寿命试验，其结果如表 8.30 所示. 在$\alpha=0.05$下，待检假设H_0，灯泡寿命服从指数分布

$$f(t)=\begin{cases}0.005e^{-0.005t}, & t\geqslant 0,\\ 0, & t<0.\end{cases}$$

表 8.30

寿命 t	[0,100)	[100,200)	[200,300)	[300,+∞)
个　数	121	78	43	58

解 待检假设$H_0:X\sim f(x)$.

当H_0为真时，可算得

$$p_i=\int_{a_{i-1}}^{a} f(t)\mathrm{d}t,\ i=1,2,3,\quad p_4=1-\sum_{i=1}^{3}p_i.$$

查表知 $$\chi^2_{0.05}(4-1)=\chi^2_{0.05}(3)=7.815.$$

因为

$n=300$，　$p_1=0.394$，　$p_2=0.239$，　$p_3=0.145$，　$p_4=0.222$，

列 χ^2 检验计算结果如表 8.31 所示，算得

$$\chi^2=\sum_{i=1}^{4}\frac{(f_i-np_i)^2}{np_i}=1.754.$$

表 8.31

区　间	f_i	p_i	np_i	f_i-np_i	$(f_i-np_i)^2/np_i$
[0,100)	121	0.394	118.2	2.8	0.0663
[100,200)	78	0.239	71.6	6.4	0.573
[200,300)	43	0.145	43.4	−0.4	0.004
[300,+∞)	58	0.222	66.6	−8.6	1.111

经比较知 $\chi^2=1.754<\chi^2_{0.05}(3)=7.815$，故接受 H_0，认为灯泡寿命服从指数分布.

例12　袋中装有8个球，其中红球数未知，在其中任取3个，记录红球的个数 x，然后放回，再任取3个，记录红球的个数然后放回. 如此重复进行了112次，其结果如表 8.32 所示，试在 $\alpha=0.05$ 下检验假设 H_0: x 服从超几何分布. 即待检假设

$$H_0: P\{x=k\}=\frac{C_5^k C_3^{3-k}}{C_8^3},\ k=0,1,2,3,$$

红球的个数为5.

表 8.32

红球个数	0	1	2	3
次　数	1	31	55	29

解　待检假设

$$H_0: P\{x=k\}=\frac{C_5^k C_3^{3-k}}{C_8^3},\ k=0,1,2,3,$$

作估计 $$\hat{p}_i=\frac{C_5^iC_3^{3-i}}{C_8^3},\ i=0,1,2,3,$$

得 $\hat{p}_0=0.066$， $\hat{p}_1=0.2678$， $\hat{p}_2=0.5371$， $\hat{p}_3=0.1785$.

列 χ^2 检验量计算表如表 8.33 所示，算得

$$\chi^2=\sum_{i=0}^{3}\frac{(f_i-np_i)^2}{np_i}=1.667.$$

查表知 $\chi^2_{0.05}(3-1)=5.991$，经比较知 $\chi^2=1.667<\chi^2_{0.05}(2)=5.991$，故接受 H_0，认为红球个数为 5.

表 8.33

x_i	f_i	p_i	np_i	f_i-np_i	$(f_i-np_i)^2/np_i$
0	1	0.0166	2 }	0	0
1	31	0.2678	30 }		
2	55	0.5371	60	−5	0.4167
3	29	0.1785	20	5	1.25

二、秩和检验法

秩和检验要注意的是，一定要分清哪类样本容量小，计算其秩. 其次，在两组临界值中，通常取犯第一类错误概率较小的一组.

例 13 对染料的某种成分进行了两次测定，分别测试了 8 瓶和 6 瓶样本，得数据如表 8.34 所示，问：这两次测定有无显著差异 $(\alpha=0.05)$？

表 8.34

第一次	2.36	3.14	7.52	3.48	2.76	5.43	6.54	7.41
第二次	4.38	4.25	6.54	3.28	7.21	6.54		

解 待检假设 $H_0:F_1(x)=F_2(x)$.

将样本值按由小到大的次序排列，确定每个样本值的秩. 因第二次样本容量较小，计算其秩. 将表 8.35 中有下画线数据的秩(属于第二个样本)相加(相同的值，取秩的平均值)，得第二样本的秩

和

$$T=4+6+7+10+10+12=49.$$

由 $\alpha=0.05, n_1=8, n_2=6$，查秩和检验表，得 $T_1=29, T_2=61$. 显然 $29<49<61$，故接受 H_0，认为两次测定无显著差异.

表 8.35

数据	2.36	2.76	3.14	3.28	3.48	4.25	4.38
秩	1	2	3	4	5	6	7
数据	5.43	6.54	6.54	6.54	7.21	7.41	7.52
秩	8	10	10	10	12	13	14

注意，当 n_1, n_2 大于 10 时，不能在表上查出. 此时可用 U 检验法，因为 $T\sim N\left(\frac{n_1}{2}(n_1+n_2+1), \frac{n_1 n_2}{12}(n_1+n_2+1)\right)$，所以其检验统计量为

$$U=\frac{T-n_1(n_1+n_2+1)/2}{\sqrt{n_1 n_2(n_1+n_2+1)/12}}\sim N(0,1),$$

拒绝域为
$$|u|>Z_{\alpha/2}.$$

本例中，$u=\frac{49-45}{7.75}=0.52<Z_{0.025}=1.96$，所以接受 H_0，可见与秩和检验结论一致.

例14 某药厂生产了一种新药，经过某医院临床试验，对服用此药的5 个病人和不服用此药的4 个病人进行跟踪调查，得到病人存活时间(单位:年)如表 8.36 所示，试确定此药是否对疾病有抑制作用($\alpha=0.05$).

表 8.36

服　药	3	4.2	3.9	5.1	4.4
不服药	2.2	0.8	0.8	1.3	

解 设总体 X 和 Y 分别为服药与不服药病人存活时间，待检假设 $H_0: \mu_1=\mu_2, H_1: \mu_1>\mu_2$.

将数据按大小排列(见表 8.37)，并取秩求秩和. 由表 8.36 算

得秩和$T=10$,查表知$T_1=12$. 显然,$10<12$,故拒绝H_0,认为此药确有抑制作用(存活时间$\mu_1>\mu_2$).

表 8.37

数据	0.8	0.8	1.3	2.2	3	3.9	4.2	4.4	5.1
秩	1	2	3	4	5	6	7	8	9

例 15 表 8.38 给出了两个地区成年居民血液中胆固醇含量的数据,试用秩和检验法检验两地区成年居民中血液胆固醇含量是否有显著差别($\alpha=0.05$).

表 8.38

地区一	403	244	253	235	319	260
地区二	403	311	269	336	259	

解 设总体X,Y分别表示地区一、地区二居民血液中胆固醇含量指标,待检假设$H_0:\mu_1=\mu_2$.

将数据按大小排列(见表 8.39),取秩,求秩和.算得

$$T=4+6+7+9+10.5=36.5.$$

表 8.39

数据	235	244	253	259	260	269	311	319	336	403	403
秩	1	2	3	4	5	6	7	8	9	10.5	10.5

查表知$T_1=20$,$T_2=40$. 显然,$20<36.5<40$,故接受H_0,认为两地区居民血液中胆固醇含量指标无显著差异.

例16 为了解甲、乙两班工人的劳动生产率,通过随机抽样取得,如表8.40 所示,试在$\alpha=0.05$下,检验甲、乙两班工人劳动生产率是否有显著差异.

表 8.40

甲班	28	33	39	40	41	42	45	46	47	
乙班	34	40	41	42	43	44	46	48	49	52

解 设总体 X 与 Y 分别表示甲、乙两班工人的劳动生产率指标，待检假设 $H_0:\mu_1=\mu_2$.

将数据按大小排列(见表 8.41)，取秩，求秩和. 算得

$$T=1+2+4+5.5+7.5+9.5+13+14.5+16=73,$$

表 8.41

甲	28	33		39	40	41	42			45	46	47			
乙			34		40	41	42	43	44		46		48	49	52
秩	1	2	3	4	5.5	7.5	9.5	11	12	13	14.5	16	17	18	19

查表知 $(n_1,n_2)=(9,10)$，得 $(T_1,T_2)=(69,111)$. 显然 $69<73<111$，故接受 H_0，认为甲、乙两班工人劳动生产率指标没有显著差异.

例17 分别从两个球队中抽查了部分队员行李的重量(单位：kg)，得数据如表 8.42 所示. 设两样本独立，且 1 队、2 队队员行李重量总体的密度至多差一个平移. 记两总体的均值分别为 μ_1,μ_2，试检验假设($\alpha=0.05$)$H_0:\mu_1=\mu_2$，$H_1:\mu_1<\mu_2$.

表 8.42

1 队	34	39	41	28	33	
2 队	36	40	35	31	39	36

解 待检假设 $H_0:\mu_1=\mu_2$，$H_1:\mu_1<\mu_2$.

将数据按大小排列(见表 8.43)，取秩，求秩和. 算得

$$T=1+3+4+8.5+11=27.5.$$

表 8.43

数据	28	31	33	34	35	36	36	39	39	40	41
秩	1	2	3	4	5	6.5	6.5	8.5	8.5	10	11

$(n_1,n_2)=(5,6)$，查表知 $(T_1,T_2)=(20,40)$. 显然，$20<27.5<40$，故接受 H_0，认为两队队员行李重量没有显著差异.

例18 A、B 两机床生产同一种零件，测得长度(单位：mm)如

表 8.44 所示，试用秩和检验法检验两机床生产的零件的长度有无显著差异($\alpha=0.05$).

表 8.44

A	20.54	27.33	29.16	21.34	24.41	20.98	29.95	17.38	21.74	31.72
B	26.27	25.09	21.85	23.39	18.41	22.60	24.64	13.62	11.84	12.77

解 设 A、B 两机床生产的零件总体分别为 X 和 Y. 待检假设 $H_0:\mu_1=\mu_2$.

将数据分别按大小排列(见表 8.45)，取秩，求秩和.

表 8.45

A				17.38		20.54	20.98	21.34	21.74	
B	11.84	12.27	13.62		18.41					21.85
秩	1	2	3	4	5	6	7	8	9	10
A			24.41				27.33	29.16	29.95	31.72
B	22.60	23.39		24.64	25.09	26.27				
秩	11	12	13	14	15	16	17	18	19	20

$n_1=n_2=10$，可任取A、B 为样本，不妨选A，则秩和$T=121$. 又$(n_1,n_2)=(10,10)$，查表知$(T_1,T_2)=(83,127)$. 显然，$83<121<127$，故接受H_0，认为两机床生产的零件长度无显著差异.

例19 为了解甲、乙两厂产品的质量，对其平均质量指标进行了抽样检测，抽得甲厂(指标μ_1)样品150 个，乙厂(指标μ_2)样品130 个，算出乙厂秩和$T=20306$. 在$\alpha=0.05$ 下，试检验假设$H_0:\mu_1=\mu_2$，$H_1:\mu_1\neq\mu_2$.

解 $n_1=150$，$n_2=130$，$T=20306$. n_1，n_2 很大，不能查表. 用检验统计量(见例 13)

$$U=\sqrt{3}\left[2T-n_1(n_1+n_2+1)\right]\Big/\sqrt{n_1n_2(n_1+n_2+1)}$$

得 $u=\sqrt{3}(40612-130\times281)\Big/\sqrt{150\times130\times281}=3.0214$.
而查表知$Z_{0.025}=1.96$，经比较知 $3.0214>1.96$，故拒绝H_0，认为$\mu_1<\mu_2$.

例 20 表8.46给出A、B两种型号的计算器充电后所能使用的时间(单位:h).设两样本独立,且数据所属总体的密度至多差一个平移,试问:能否认为A型计算器充电后使用时间比B型的计算器长($\alpha=0.01$)?

表 8.46

A	5.5	5.6	6.3	4.6	5.3	5.0	6.2	5.8	5.1	5.2	5.9	
B	3.8	4.3	4.2	4.0	4.9	4.5	5.2	4.8	4.5	3.9	3.7	4.6

解 待检假设

$$H_0:\mu_A=\mu_B,\quad H_1:\mu_A>\mu_B.$$

将数据分别按大小排列(见表8.47),取秩,求秩和.算得

$$T=9.5+13+14+15.5+17+18+19+20+21+22+23=192.$$

表 8.47

数据	3.7	3.8	3.9	4.0	4.2	4.3	4.5	4.5	4.6	4.6	4.8	4.9
秩	1	2	3	4	5	6	7.5	7.5	9.5	9.5	11	12
数据	5.0	5.1	5.2	5.2	5.3	5.5	5.6	5.8	5.9	6.2	6.3	
秩	13	14	15.5	15.5	17	18	19	20	21	22	23	

又 $n_1(n_1+n_2+1)/2=132,\quad n_1n_2(n_1+n_2+1)/12=264.$

计算检验统计量U的值(因$n_1,n_2>10$),得

$$u=\frac{192-132}{\sqrt{264}}=3.69.$$

查表知$Z_{0.01}=2.33$,经比较知$u=3.69>Z_{0.01}=2.33$,故拒绝H_0,认为A型计算器充电后使用时间比B型计算器使用时间长.

硕士研究生入学试题分析

一、本章考试要求

1. 理解显著性检验的基本思想,掌握假设检验的基本步骤,了解假设检验可能产生的两类错误.

2. 掌握单个及两个正态总体的均值和方差的假设检验.

二、本章的重点内容

假设检验(单个或两个正态总体的均值和方差),确定检验统计量并进行假设检验.

1. 设 $X_1,X_2,\cdots,X_n$ 是来自正态总体 $N(\mu,\sigma^2)$ 的简单随机样本,其中参数 μ 和 σ^2 未知. 记 $\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i, Q^2=\sum_{i=1}^{n}(X_i-\overline{X})^2$,则假设 $H_0:\mu=0$ 的 T 检验使用统计量________. (1995 年四)

解 $T=(\overline{X}-\mu)/(S/\sqrt{n})$,化为

$$T=\frac{\bar{x}}{S/\sqrt{n}}=\bar{x}\sqrt{n}\Big/\sqrt{\sum_{i=1}^{n}(X_i-\overline{X})^2/(n-1)}$$

$$=\bar{x}\sqrt{n(n-1)}\Big/\sqrt{\sum_{i=1}^{n}(X_i-\overline{X})^2}=\bar{x}\sqrt{n(n-1)}/Q.$$

2. 设某次考试的考生成绩服从正态分布,从中随机地抽取 36 位考生的成绩,算得平均成绩为 66.5 分,标准差为 15 分. 问:在显著性水平 0.05 下,是否可以认为这次考试全体考生的平均成绩为 70 分? 并给出检验过程. 表 8.48 是 t 分布表 $P\{t(n)\leqslant t_p(n)\}=p$.

表 8.48

$t_p(n)$ \ p / n	0.95	0.975
35	1.6896	2.0301
36	1.6883	2.0281

(1998 年一)

解 设考生成绩 $X\sim N(\mu,\sigma^2)$,样本容量 $n=36,\bar{x}=66.5,S=15$,待检假设 $H_0:\mu=70$.

因为 σ^2 未知,检验 μ 用 T 检验法,得

$$|t|=|\bar{x}-70|\times\sqrt{n}/S=|66.5-70|\times 6/15=1.4.$$

查表知 $t_{0.025}(36-1)=t_{0.025}(35)=2.0301$,经比较知 $|t|=1.4<t_{0.025}(35)=2.0301$,故接受 H_0,认为平均成绩是 70 分.

第九章　方差分析与回归分析

第一节　方 差 分 析

主 要 内 容

方差分析(analysis of variance)是对试验的结果进行分析，鉴别不同因素对试验结果影响大小的一种有效方法，是英国统计学家费舍尔(R. A. Fisher)最先提出并使用的.

一、单因素试验的方差分析

影响试验指标的条件称为试验的因素，只有一个因素在改变的试验称为单因素试验.

因素所处的状态称为因素的水平. 设试验的因素有 s 个水平 $A_1,A_2,\cdots,A_s$，在各个水平 A_j $(j=1,2,\cdots,s)$下的样本 $x_{1j},x_{2j},\cdots,x_{nj}$ 来自正态总体 $X\sim N(\mu_j,\sigma^2)$，且相互独立，其中 μ_j,σ^2 未知. 在正态总体和方差齐性(σ^2 相同)前提下，提出待检假设

H_0:　　$\mu_1=\cdots=\mu_s=\mu_0$，　H_1: μ_i 不全相等.

样本观察值如表 9.1 所示.

记 x_{ij} 为第 j 个水平下第 i 次试验的结果.

$$n=n_1+n_2+\cdots+n_s,$$

表中

$$\bar{x}_{\cdot j}=\frac{1}{n_j}\sum_{i=1}^{n_j}x_{ij}\quad(j=1,2,\cdots,s),$$

$$T_{\cdot j}=\sum_{i=1}^{n_j}x_{ij}=n_j\bar{x}_{\cdot j}\quad(j=1,2,\cdots,s),$$

表 9.1

水平 / 观察值	A_1	A_2	…	A_s
	x_{11}	x_{12}	…	x_{1s}
	x_{21}	x_{22}	…	x_{2s}
	⋮	⋮		⋮
	$x_{n_1 1}$	$x_{n_2 2}$	…	$x_{n_s s}$
样本总和	$T_{\cdot 1}$	$T_{\cdot 2}$	…	$T_{\cdot s}$
样本均值	$\bar{x}_{\cdot 1}$	$\bar{x}_{\cdot 2}$	…	$\bar{x}_{\cdot s}$
总体均值	μ_1	μ_2	…	μ_s

$\bar{x}=\dfrac{1}{n}\sum\limits_{j=1}^{s}\sum\limits_{i=1}^{j}x_{ij}$，称为数据总平均.

$T=\sum\limits_{j=1}^{s}\sum\limits_{i=1}^{n_j}x_{ij}=n\bar{x}$，$\mu=\dfrac{1}{n}\sum\limits_{j=1}^{s}n_j\mu_j$，称为总平均.

$S_T=\sum\limits_{j=1}^{s}\sum\limits_{i=1}^{n_j}(x_{ij}-\bar{x})^2$ 称为总变差平方和.

$S_A=\sum\limits_{j=1}^{s}\sum\limits_{i=1}^{n_j}(\bar{x}_{\cdot j}-\bar{x})^2=\sum\limits_{j=1}^{s}n_j(\bar{x}_{\cdot j}-\bar{x})^2$ 称为效应平方和.

$S_E=\sum\limits_{j=1}^{s}\sum\limits_{i=1}^{n_j}(x_{ij}-\bar{x}_{\cdot j})^2$ 称为误差平方和.

存在关系
$$S_T=S_A+S_E.$$
由于 $\dfrac{S_T}{\sigma^2}\sim\chi^2(n-1)$，$\dfrac{S_E}{\sigma^2}\sim\chi^2(n-s)$，$\dfrac{S_A}{\sigma^2}\sim\chi^2(s-1)$，
$$E(S_E)=(n-s)\sigma^2,\quad E(S_A)=(s-1)\sigma^2.$$
S_A 与 S_E 相互独立. 选择统计量
$$F=\frac{S_A/(s-1)}{S_E/(n-s)}\sim F(s-1,n-s).$$
检验的拒绝域为 $F\geqslant F_\alpha(s-1,n-s)$（$\alpha$ 为显著性水平）.

单因素试验方差分析表（见表 9.2）反映了方差分析的结果.

表 9.2

方差来源	平方和	自由度	均　方	$F_{比}$	结　论
因素 A	S_A	$s-1$	$\overline{S}_A=S_A/(s-1)$	$F=\overline{S}_A/\overline{S}_E$	
误差	S_E	$n-s$	$\overline{S}_E=S_E(n-s)$		
总和	S_T	$n-1$			

二、双因素试验的方差分析

有两个因素 A,B 作用于试验的某一指标. 因素 A 有 r 个水平 $A_1,A_2,\cdots,A_r$，因素 B 有 s 个水平 $B_1,B_2,\cdots,B_s$，对因素 A,B 的水平的每对组合 (A_i,B_j)，$i=1,2,\cdots,r$ 且 $j=1,2,\cdots,s$ 做试验. 按试验次数分为无重复试验和等重复试验.

1. 双因素无重复试验的方差分析

对因素 A,B 的水平的每对组合 (A_i,B_j) 只做一次试验的方差分析称双因素无重复试验的方差分析. 将试验数据列于表 9.3.

表 9.3

因素 B / 因素 A	B_1	B_2	$\cdots$	B_s	$\overline{x}_{i\cdot}$
A_1	x_{11}	x_{12}	$\cdots$	x_{1s}	$\overline{x}_{1\cdot}$
A_2	x_{21}	x_{22}	$\cdots$	x_{2s}	$\overline{x}_{2\cdot}$
$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$
A_r	x_{r1}	x_{r2}	$\cdots$	x_{rs}	$\overline{x}_{r\cdot}$
$\overline{x}_{\cdot j}$	$\overline{x}_{\cdot 1}$	$\overline{x}_{\cdot 2}$	$\cdots$	$\overline{x}_{\cdot s}$	$\overline{x}$

设 $x_{ij}\sim N(\mu_{ij},\sigma^2)$，是从正态总体 $N(\mu_{ij},\sigma^2)$ 中抽得的容量为 1 的样本，且相互独立. 其中

$$\mu_{ij}=\mu+\alpha_i+\beta_j,\ i=1,2,\cdots,r \text{ 且 } j=1,2,\cdots,s,$$

$$\sum_{i=1}^{r}\alpha_i=0,\quad \sum_{j=1}^{s}\beta_j=0.$$

α_i 称为因素 A 在水平 A_i 的效应，β_j 称为因素 B 在水平 B_j 的效应.

作假设 $H_{01}:\alpha_1=\alpha_2=\cdots=\alpha_r=0$，又假设 $H_{02}:\beta_1=\beta_2=\cdots=\beta_s=0$，记

$$\bar{x}=\frac{1}{rs}\sum_{i=1}^{r}\sum_{j=1}^{s}x_{ij},\quad \bar{x}_{i\cdot}=\frac{1}{s}\sum_{j=1}^{s}x_{ij},\ i=1,2,\cdots,r,$$

$$\bar{x}_{\cdot j}=\frac{1}{r}\sum_{i=1}^{r}x_{ij},\ j=1,2,\cdots,s,$$

则

$$S_T=\sum_{i=1}^{r}\sum_{j=1}^{s}(x_{ij}-\bar{x})^2$$

称为总变差平方和；

$$S_A=s\sum_{i=1}^{r}(\bar{x}_{i\cdot}-\bar{x})^2,\quad S_B=r\sum_{j=1}^{s}(\bar{x}_{\cdot j}-\bar{x})^2$$

称为因素 A,B 的效应平方和；

$$S_E=\sum_{i=1}^{r}\sum_{j=1}^{s}(x_{ij}-\bar{x}_{i\cdot}-\bar{x}_{\cdot j}+\bar{x})^2$$

称为因素 A,B 的误差平方和.

当 H_{01} 为真时，

$$F_A=\frac{S_A/(r-1)}{S_E/[(r-1)(s-1)]}\sim F((r-1),(s-1)(r-1)),$$

在显著性水平 α 下，拒绝域形式为

$$F_A\geqslant F_\alpha((r-1),(s-1)(r-1)).$$

当 H_{02} 为真时，

$$F_B=\frac{S_B/(s-1)}{S_E/[(r-1)(s-1)]}\sim F((s-1),(r-1)(s-1)),$$

在显著性水平 α 下，拒绝域形式为

$$F_B\geqslant F_\alpha((s-1),(r-1)(s-1)).$$

表 9.4 是双因素无重复试验方差分析表.

表 9.4

方差来源	平方和	自由度	均　方	$F_{比}$	结论
因素 A	S_A	$r-1$	$\bar{S}_A=S_A/(r-1)$	$F_A=\bar{S}_A/\bar{S}_E$	
因素 B	S_B	$s-1$	$\bar{S}_B=S_B/(s-1)$	$F_B=\bar{S}_B/\bar{S}_E$	
误差	S_E	$(r-1)(s-1)$	$\bar{S}_E=S_E/(r-1)(s-1)$		
总和	S_T	$rs-1$			

2. 双因素等重复试验的方差分析

对因素 A,B 的水平的每对组合 (A_i,B_j) 都作相同次数重复试验的方差分析称为双因素等重复试验的方差分析，数据如表 9.5 所示.

表 9.5

因素B / 因素A	B_1	B_2	$\cdots$	B_s	$\bar{x}_{i\cdot\cdot}$
A_1	$\bar{x}_{11\cdot}$	$\bar{x}_{12\cdot}$	$\cdots$	$\bar{x}_{1s\cdot}$	$\bar{x}_{1\cdot\cdot}$
A_2	$\bar{x}_{21\cdot}$	$\bar{x}_{22\cdot}$	$\cdots$	$\bar{x}_{2s\cdot}$	$\bar{x}_{2\cdot\cdot}$
$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$
A_r	$\bar{x}_{r1\cdot}$	$\bar{x}_{r2\cdot}$	$\cdots$	$\bar{x}_{rs\cdot}$	$\bar{x}_{r\cdot\cdot}$
$\bar{x}_{\cdot j\cdot}$	$\bar{x}_{\cdot 1\cdot}$	$\bar{x}_{\cdot 2\cdot}$	$\cdots$	$\bar{x}_{\cdot s\cdot}$	—

$$\bar{x}=\frac{1}{rst}\sum_{i=1}^{r}\sum_{j=1}^{s}\sum_{k=1}^{t}x_{ijk},$$

$$\bar{x}_{ij\cdot}=\frac{1}{t}\sum_{k=1}^{t}x_{ijk},\quad i=1,2,\cdots,r \text{ 且 } j=1,2,\cdots,s,$$

记
$$\bar{x}_{i\cdot\cdot}=\frac{1}{st}\sum_{j=1}^{s}\sum_{k=1}^{t}x_{ijk},\quad i=1,2,\cdots,r,$$

$$\bar{x}_{\cdot j\cdot}=\frac{1}{rt}\sum_{i=1}^{r}\sum_{k=1}^{t}x_{ijk},\quad j=1,2,\cdots,s.$$

而总平方和及其分解式为

$$S_T=\sum_{i=1}^{r}\sum_{j=1}^{s}\sum_{k=1}^{t}(x_{ijk}-\bar{x})^2,$$

$$S_T=S_A+S_B+S_{AB}+S_E.$$

其中
$$S_E=\sum_{i=1}^{r}\sum_{j=1}^{s}\sum_{k=1}^{t}(x_{ijk}-\bar{x}_{ij\cdot})^2,$$

$$S_A=st\sum_{i=1}^{r}(\bar{x}_{i\cdot\cdot}-\bar{x})^2,$$

$$S_B=rt\sum_{j=1}^{s}(\bar{x}_{\cdot j\cdot}-\bar{x})^2,$$

$$S_{AB}=t\sum_{i=1}^{r}\sum_{j=1}^{s}(\bar{x}_{ij\cdot}-\bar{x}_{i\cdot\cdot}-\bar{x}_{\cdot j\cdot}+\bar{x})^2.$$

S_{AB}称为 A,B 的交互效应平方和.

待检假设

$H_{01}:\alpha_1=\alpha_2=\cdots=\alpha_r=0$ （α_i 是水平 A_i 的效应），

$H_{02}:\beta_1=\beta_2=\cdots=\beta_s=0$ （β_j 是水平 β_j 的效应），

$H_{03}:\gamma_{11}=\gamma_{12}=\cdots=\gamma_{rs}=0$ （γ_{ij}是 A_i 和 B_j 的交互效应）.

当H_{01}为真时，

$$F_A=\frac{S_A/(r-1)}{S_E/[rs(t-1)]}\sim F((r-1),rs(t-1)),$$

拒绝域为 $F_A\geqslant F_\alpha((r-1),rs(t-1))$.

当H_{02}为真时，

$$F_B=\frac{S_B/(s-1)}{S_E/[rs(t-1)]}\sim F((s-1),rs(t-1)),$$

拒绝域为 $F_B\geqslant F_\alpha((s-1),rs(t-1))$.

当H_{03}为真时，

$$F_{AB}=\frac{S_{AB}/[(r-1)(s-1)]}{S_E/[rs(t-1)]}\sim F((s-1)(r-1),rs(t-1)).$$

分析结果列成双因素等重复试验方差分析表（见表9.6）.

表 9.6

方差来源	平方和	自由度	均　方	$F_{比}$	结论
因素 A	S_A	$r-1$	$\overline{S}_A=S_A/(r-1)$	$F_A=\overline{S}_A/\overline{S}_E$	
因素 B	S_B	$s-1$	$\overline{S}_B=S_B/(s-1)$	$F_B=\overline{S}_B/\overline{S}_E$	
交互作用	S_{AB}	$(r-1)(s-1)$	$\overline{S}_{AB}=S_{AB}/(r-1)(s-1)$	$F_{AB}=\overline{S}_{AB}/\overline{S}_E$	
误差	S_E	$rs(t-1)$	$\overline{S}_E=S_E/rs(t-1)$		
总和	S_T	$rst-1$			

疑 难 解 析

1. 怎样区分所讨论的问题是方差分析还是回归分析？

答 方差分析与回归分析都是考察所研究的某一指标与试验因素（条件）的关系的. 方差分析考察的是因素对指标的影响是否

显著，而回归分析考察的是因素的取值与指标的取值存在一种什么样的相关关系.

因素可以分为两大类，一类是属性的，一类是数量的. 属性的因素一般无数量大小可言，只是性质的不同，如种子的品种、机器的型号、材料的品质、加工的工艺等等. 数量的因素可以在一定范围内取值，如人的身高、体重，试验的温度，产量，产品的合格率等等. 也有本来是数量而属性化的，如施肥量可以是某个数量，但有时将它局限在某些范围内而分为高、中、低几个层次，就属性化了.

当所考虑问题的因素是属性的时，问题属于方差分析的范畴；当所考虑的因素是数量的时，问题属于回归分析的范畴.

2. 方差分析的依据是什么？

答 方差分析的种类很多. 在不同类型的方差分析中，因素可以增加或减少，数据结构可以发生变化. 但是以下三个重要的假定是不变的.

(1) 正态性假定　有了正态性假定后，数据 x_{ij} 认为取自 $N(\mu,\sigma^2)$，由此求得的各种离差平方和(如比值 S_T/σ^2)$\sim\chi^2$ 分布，从而定义 F 分布函数. 没有正态假定，就没有 χ^2 分布，也没有 F 分布与统计推断.

(2) 方差齐性假定　假定数据 x_{ij} 来自方差为 σ^2 的正态总体，只有这样才能在相同的条件(σ^2 相等)下来分析问题. 考察指标的变化，才可以建立统计假设 H_0，才有方差分析检验.

(3) 线性假定　线性假定指数据 x_{ij} 的取得仅通过线性运算，这样才可以把数据 x_{ij} 当线性模型处理，也才可以施行方差分析方法.

在大数定律和中心极限定理下，正态性假设是易于确立的. 数据的线性假设也符合实际，易于成立，但是，方差齐性假设不易确立. 例如对于二项分布来说，其样本的方差 $S^2=\dfrac{p(1-p)}{n}$ 随 p 而变化，对于不同组数据，很难保持方差齐性，所以常常用数据的变换来实现. 由于其计算比较复杂，本书不加叙述.

在方差分析中，三个假定缺一不可，否则方差分析就失去了依据. 在疑难解析问题4 中将提到，方差分析可认为是 t 检验的发展，而对两个总体均值差在方差未知时的检验正是在正态总体、方差齐性($\sigma_1^2=\sigma_2^2=\sigma^2$)、线性数据的假定下进行的，可见两者十分一致.

3. 方差分析中所考虑因素的多少是怎样确定的?

答 在实际问题中，影响指标的因素往往是很多的，这就需要根据问题的性质，对问题的了解和研究的规模确定因素的取舍，即只研究其中哪几个(或一个)因素，将因素分几个水平，怎样区分水平. 为了研究方便且突出结论，一般只让一个或两个因素变化，而把其它因素固定起来. 这样观察和分析指标的变化和因素的影响，就得到单因素和双因素方差分析.

4. 方差分析与 t 检验有什么关系?

答 方差分析是由 t 检验发展来的，它们都是检验几个总体的平均数是否来自同一正态总体的可信程度的.

当因素的水平等于 2 时，方差分析检验与 T 检验是一致的. 当因素的水平(单因素)等于 2 时，F 统计量的第一自由度是 1，F 分布表中的 F 值与 t 分布表中的 t 值存在平方根关系，即 $t_{\alpha/2}(1)=\sqrt{F_\alpha(1,1)}$. 这说明 t 函数的平方是一个 F 函数. 但是 t 检验只适合水平等于 2 的情形，而方差分析适合水平大于等于 2 的情形.

5. 方差分析与 F 检验有何关系?

答 方差分析与 F 检验的关系在前面已经提到，方差分析是检验几个总体的平均数来自同一正态总体的可信程度，而 F 检验法是检验两个总体的方差来自同一正态总体的可信程度，两者的出发点是不同的.

在方差分析中，规定了 $S_A/(s-1)$ 是第一样本作分子，$S_E/(n-s)$ 是第二样本作分母. 而 F 检验事先没有规定哪个样本为分子或分母，而是在算出数值后，通常以数值大的那个作第一样本写在分子上，这又是一个重要差别.

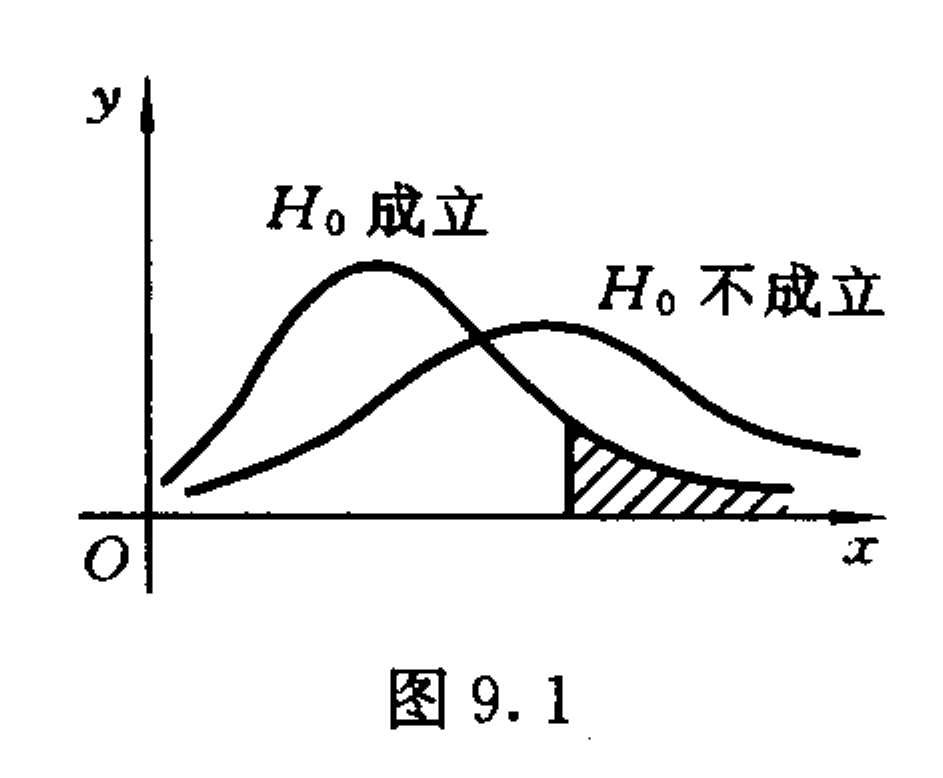


图 9.1

在方差分析中，当 H_0 成立时，$[S_A/(s-1)]/[S_E/(n-s)]$服从 F 分布；当 H_0 不成立时，分布要偏右些（见图 9.1）. 因而，在方差分析时，通常采用单侧检验. 在 F 检验中，H_0 可信成立时，两样本的方差比服从 F 分布；H_0 不成立时，分布可能偏左也可能偏右.

方法、技巧与典型例题分析

一、单因素方差分析

在对具体问题进行方差分析时，首先，应区分实际问题是单因素方差分析还是双因素方差分析，其中主要考虑哪些是可以固定的因素，哪些是变化的因素；其次，从数据来观察，看其是重复试验还是非重复试验.

在计算过程中，要注意使用简化公式.

单因素方差分析的简化公式是：

记 $\quad T_{\cdot j}=\sum_{i=1}^{n_j}x_{ij},\ j=1,2,\cdots,s,\quad T=\sum_{j=1}^{s}\sum_{i=1}^{n_j}x_{ij},$

即有
$$\begin{cases}S_T=\sum_{j=1}^{s}\sum_{i=1}^{n_j}x_{ij}^2-n\bar{x}^2=\sum_{j=1}^{s}\sum_{i=1}^{n_j}x_{ij}^2-\dfrac{T^2}{n},\\ S_A=\sum_{j=1}^{s}n_j\bar{x}_{\cdot j}^2-n\bar{x}^2=\sum_{j=1}^{s}\dfrac{T_{\cdot j}^2}{n_j}-\dfrac{T^2}{n},\\ S_E=S_T-S_A.\end{cases}$$

如果 s 个样本的容量都相同，即 $n_1=n_2=\cdots=n_s=n_0$，则称之为等重复试验，否则称之为不等重复试验. 若是等重复试验，有

$$S_A=\frac{1}{n_0}\sum_{j=1}^{s}T_{\cdot j}^2-\frac{T^2}{n}.$$

例 1 对于某种作物进行 5 种不同肥料的耕作试验，每种肥料

做 4 次试验，试验的收获量（单位：kg）如表 9.7 所示，问：不同的肥料对收获量有无显著的影响（$\alpha=0.05$）？

表 9.7

		试验批号			
		1	2	3	4
肥料	A_1	67	67	45	52
	A_2	98	96	91	66
	A_3	60	69	50	35
	A_4	79	64	81	70
	A_5	90	70	79	88

解 分别以 $\mu_1,\mu_2,\cdots,\mu_5$ 表示 5 种肥料下收获量总体的均值，待检假设 $H_0:\mu_1=\mu_2=\mu_3=\mu_4=\mu_5$. 经计算得

$$s=5,\quad n_0=4,\quad n=20,\quad \alpha=0.05,$$

$$S_T=\sum_{j=1}^{5}\sum_{i=1}^{4}x_{ij}^2-\frac{1}{20}\Big(\sum_{j=1}^{5}\sum_{i=1}^{4}x_{ij}\Big)^2$$

$$=106033-\frac{1}{20}\times1417^2=5638.55,$$

$$S_A=\frac{1}{4}\sum_{j=1}^{5}\Big(\sum_{i=1}^{4}x_{ij}\Big)^2-\frac{1}{20}\Big(\sum_{j=1}^{5}\sum_{i=1}^{4}x_{ij}\Big)^2$$

$$=\frac{1}{4}\times415723-\frac{1}{20}\times1417^2=3536.30.$$

$$S_E=S_T-S_A=2102.25.$$

而 $n-1=19$，$n-s=15$，$s-1=4$，方差分析结果如表 9.8 所示. 由 $F_{0.05}(4,15)=3.06<F_{比}=6.308$，故拒绝 H_0，认为不同肥料对收获量的影响是高度显著的.

表 9.8

方差来源	平方和	自由度	均方	$F_{比}$	结论
因素	3536.30	4	884.075		显著
误差	2102.25	15	140.15	6.308	
总和	5638.55	19			

例2 有一些棉布用不同的印染工艺处理,然后进行缩水率试验.假设采用了5种不同工艺,每种工艺处理4块布样,测得缩水率(%)如表9.9所示.

若布的缩水率服从正态分布,不同工艺下处理布的缩水率方差相等,试考察不同工艺对布的缩水率有无显著影响($\alpha=0.05$).

表9.9

		试验批号			
		1	2	3	4
因素	A_1	4.3	7.8	3.2	6.5
	A_2	6.1	7.3	4.2	4.1
	A_3	4.3	8.7	7.2	10.1
	A_4	6.5	8.3	8.6	8.2
	A_5	9.5	8.8	11.4	7.8

解 待检假设$H_0: \mu_1=\mu_2=\mu_3=\mu_4=\mu_5$. μ_i为不同工艺下缩水率总体的均值.为简便起见,将每一数据减去7.4,再除以0.1,列出方差计算表(变换后数据仍记为x_{ij},平方和仍分别为S_T, S_A, S_B).

$$\sum_{j=1}^{5}\sum_{i=1}^{4}x_{ij}^2=9591, \quad T=51, \quad T^2=2601, \quad \sum_{j=1}^{5}T_{\cdot j}^2=19015.$$

$$S_T=\sum_{j=1}^{5}\sum_{i=1}^{4}x_{ij}^2-\frac{1}{20}T^2=9460.95,$$

$$S_A=\frac{1}{4}\sum_{j=1}^{5}T_{\cdot j}^2-\frac{1}{20}T^2=4623.70,$$

$$S_E=S_T-S_A=4837.25.$$

方差分析结果如表9.10所示.因为$F_{0.05}(4,15)=3.06<F_{比}=3.58$,故拒绝$H_0$,认为不同印染工艺对布的缩水率有显著影响.

表9.10

方差来源	平方和	自由度	均方	$F_{比}$	结论
因素	4623.70	4	1155.925	3.58	显著
误差	4837.25	15	332.483		
总和	9460.95	19			

例3 有三台机器,生产同一种规格的铝合金薄板.测量三台

机器所生产的薄板厚度(单位:mm),结果如表 9.11 所示. 试考察机器对薄板厚度有无显著的影响($\alpha=0.05$).

表 9.11

机器 1	机器 2	机器 3
0.236	0.257	0.258
0.238	0.253	0.264
0.248	0.255	0.259
0.245	0.254	0.267
0.243	0.261	0.262

解 待检假设 $H_0:\mu_1=\mu_2=\mu_3$. μ_i 是各台机器生产的薄板总体的均值. 经计算得

$$s=3,\quad n_1=n_2=n_3=5,\quad n=15,$$

$$\sum_{j=1}^{3}\sum_{i=1}^{5}x_{ij}^2=0.963912,\quad T=3.8,\quad \sum_{j=1}^{3}T_{\cdot j}^2=4.8102.$$

$$S_T=\sum_{j=1}^{3}\sum_{i=1}^{5}x_{ij}^2-\frac{1}{15}T^2=0.001245,$$

$$S_A=\frac{1}{5}\sum_{j=1}^{3}T_{\cdot j}^2-\frac{1}{15}T^2=0.001053,$$

$$S_E=S_T-S_A=0.000192.$$

方差分析结果如表 9.12 所示. 因为 $F_{0.05}(2,12)=3.89<F_{比}=32.92$,故拒绝 H_0,认为各台机器生产的薄板厚度有显著差异.

表 9.12

方差来源	平方和	自由度	均方	$F_{比}$	结论
因素	0.00105	2	0.005266	32.92	显著
误差	0.000192	12	0.000016		
总和	0.001245	14			

在进行方差分析时,还常要对未知参数进行估计. 下面写出常

用的几个估计：

(1) $\hat{\sigma}^2=\dfrac{S_E}{n-s}$是$\sigma^2$ 的无偏估计.

(2) $\hat{\mu}=\bar{x}$，$\hat{\mu}_j=\bar{x}_{\cdot j}$分别是$\mu$，$\mu_j$ 的无偏估计.

(3) $\hat{\sigma}_j=\bar{x}_{\cdot j}-\bar{x}$ 是δ_j 的无偏估计，且$\sum\limits_{j=1}^{s}n_j\delta_j=0$.

(4) 两总体$N(\mu_j,\sigma^2)$与$N(\mu_k,\sigma^2)$的均值差$\mu_j-\mu_k$ 的置信度为$1-\alpha$ 的置信区间为

$$\left(\bar{x}_{\cdot j}-\bar{x}_{\cdot k}\mp t_{\alpha/2}(n-s)\sqrt{S_E(1/n_j+1/n_k)}\right).$$

例 4 求上例中未知参数σ^2，μ_j，δ_j 的点估计及均值差的置信度为 0.95 的置信区间.

解 $\hat{\sigma}^2=\dfrac{S_E}{n-s}=\dfrac{0.000192}{15-3}=0.000016$，

$\hat{\mu}_1=\bar{x}_{\cdot 1}=0.242$， $\hat{\mu}_2=\bar{x}_{\cdot 2}=0.256$， $\hat{\mu}_3=\bar{x}_{\cdot 3}=0.262$，

$$\hat{\mu}=\bar{x}=0.253,\quad \hat{\delta}_1=\bar{x}_{\cdot 1}-\bar{x}=-0.011,$$

$$\hat{\delta}_2=\bar{x}_{\cdot 2}-\bar{x}=0.003,\quad \hat{\delta}_3=\bar{x}_{\cdot 3}-\bar{x}=0.009.$$

又由
$$t_{0.025}(15-3)=2.1788,$$

$$\sqrt{S_E(1/n_j+1/n_k)}=\sqrt{16\times10^{-6}\times\frac{2}{5}}=1.265\times10^{-3},$$

知
$$t_{0.025}(12)\sqrt{S_E(1/n_j+1/n_k)}=0.0055.$$

故$\mu_1-\mu_2$，$\mu_1-\mu_3$ 及$\mu_2-\mu_3$ 的置信度为 0.95 的置信区间分别为

$$(0.242-0.256\mp0.0055)=(-0.0195,-0.0085),$$

$$(0.242-0.262\mp0.0055)=(-0.0255,-0.0145),$$

$$(0.256-0.262\mp0.0055)=(-0.0115,-0.0005).$$

例 5 有同一型号的电池三批，它们分别是 A、B、C 三个工厂生产的. 现各随机抽取 5 只电池，经试验测得其寿命(单位：h)如表 9.13 所示. 试在显著性水平$\alpha=0.05$ 下检验电池的平均寿命有无显著差异，并求$\mu_A-\mu_B$，$\mu_A-\mu_C$，$\mu_B-\mu_C$ 的置信度为 0.95 的置信区

间. 设各厂电池寿命服从同方差的正态分布.

表 9.13

A厂	40	48	38	42	45
B厂	26	34	30	28	32
C厂	39	40	43	50	50

解 以 μ_A, μ_B, μ_C 记各厂生产电池的平均寿命，待检假设 $H_0: \mu_A=\mu_B=\mu_C$. 计算得

$$s=3, \quad n_1=n_2=n_3=5, \quad n=15,$$

$$S_T=\sum_{j=1}^{3}\sum_{i=1}^{5}x_{ij}^2-\frac{1}{15}T^2=832,$$

$$S_A=\frac{1}{5}\sum_{j=1}^{3}T_{\cdot j}^2-\frac{1}{15}T^2=615.6,$$

$$S_E=S_T-S_A=216.4,$$

$$\overline{S}_A=\frac{S_A}{2}=307.8, \quad \overline{S}_E=\frac{S_E}{12}=18.03,$$

$$F_{比}=\overline{S}_A/\overline{S}_E=17.07.$$

方差分析结果如表 9.14 所示. 因为 $F_{0.05}(2,12)=3.89<F_{比}=17.07$，故拒绝 H_0，认为各厂生产的电池寿命差异显著.

表 9.14

方差来源	平方和	自由度	均方	$F_{比}$	结论
因素	615.6	2	307.8	17.07	显著
误差	216.4	12	18.03		
总和	832	14			

作出如下估计：

$$\hat{\mu}_A=\bar{x}_A=42.6, \quad \hat{\mu}_B=\bar{x}_B=30, \quad \hat{\mu}_C=\bar{x}_C=44.4,$$

$$t_{0.025}(14-2)\sqrt{\overline{S}_E(1/n_j+1/n_k)}$$

$$=2.1788\sqrt{18.03\times 2/5}=5.85,$$

$\mu_A-\mu_B, \mu_A-\mu_C, \mu_B-\mu_C$ 的置信度为 0.95 的置信区间分别为

(6.75,18.45),(−7.65,4.05),(−20.25,−8.55).

例6 某年级有三个班进行了一次数学考试,从各班随机抽取部分学生,记录其数学成绩如表 9.15 所示. 试在显著性水平 $\alpha=0.05$下检验各班成绩有无显著差异. 设各总体是正态总体,且方差相等.

表 9.15

1班				2班					3班				
73	66	89	60	88	77	78	31	48	87	68	41	79	59
82	45	93	80	78	91	62	51	76	71	56	68	91	53
36	77	43	73	85	96	74	80	56		79	71	15	

解 以$\mu_i(i=1,2,3)$记第i 班的平均成绩,待检假设$H_0:\mu_1=\mu_2=\mu_3$. 计算得

$$s=3,\quad n_1=12,\quad n_2=15,\quad n_3=13,\quad n=40,$$

$$S_T=\sum_{j=1}^{3}\sum_{i=1}^{n_j}x_{ij}^2-\frac{1}{40}T^2=13685.1,$$

$$S_A=\sum_{j=1}^{3}\frac{1}{n_j}T_{\cdot j}^2-\frac{1}{40}T^2=335.35,$$

$$S_E=S_T-S_A=13349.75.$$

方差分析结果如表 9.16 所示. 因为 $F_{0.05}(2,31)=3.32>F_{比}=0.465$,故接受$H_0$,认为各班成绩无显著差异.

表 9.16

方差来源	平方和	自由度	均 方	$F_{比}$	结论
因素	335.35	2	167.68	0.465	不显著
误差	13349.75	31	360.80		
总和	13685.1	39			

例7 对单因素方差分析问题,求σ^2 的置信度为$1-\alpha$ 的置信区间.

解 因为不论H_0是否为真,$\hat{\sigma}^2=\dfrac{S_E}{n-s}$都是$\sigma^2$ 的无偏估计,$\dfrac{S_E}{\sigma^2}$

$\sim\chi^2(n-s)$，这里 s 是因素的水平个数，所以

$$P\left\{\chi^2_{1-\alpha/2}(n-s)<\frac{S_E}{\sigma^2}\leqslant\chi^2_{\alpha/2}(n-s)\right\}=1-\alpha,$$

得 σ^2 的置信度为 $1-\alpha$ 的置信区间为

$$\left(\frac{S_E}{\chi^2_{\alpha/2}(n-s)},\frac{S_E}{\chi^2_{1-\alpha/2}(n-s)}\right).$$

例 8 有四种类型的用于计算器电路的响应时间(单位:ms)如表 9.17 所示. 设响应时间总体均为正态总体，且各总体方差相同，各样本相互独立，试在 $\alpha=0.05$ 下检验各类型电路的响应时间有无显著差异.

表 9.17

类型 Ⅰ	类型 Ⅱ	类型 Ⅲ	类型 Ⅳ
19 15 22	20 40 21	16 17 15	18 22 19
20 18	33 27	18 26	

解 以 μ_i $(i=1,2,\cdots,4)$ 记四个类型电路响应时间总体的平均值. 待检假设 $H_0:\mu_1=\mu_2=\mu_3=\mu_4$. 计算得

$$n=18,\quad s=4,\quad n_1=n_2=n_3=5,\quad n_4=3,$$

$$S_T=\sum_{j=1}^{4}\sum_{i=1}^{n_j}x_{ij}^2-\frac{1}{18}T^2$$

$$=8992-\frac{1}{18}\times386^2=714.44,$$

$$S_A=\sum_{j=1}^{4}\frac{1}{n_j}T_{\cdot j}^2-\frac{1}{18}T^2$$

$$=\left[\frac{1}{5}\times(94^2+141^2+92^2)+\frac{1}{3}\times59^2\right]-\frac{1}{18}\times386^2$$

$$=318.98,$$

$$S_E=S_T-S_A=395.46.$$

方差分析结果如表 9.18 所示. 因为 $F_{0.05}(3,14)=3.34<F_{比}=3.76$，故拒绝 H_0，认为各类型电路响应时间有显著差异.

表 9.18

方差来源	平方和	自由度	均方	$F_{比}$	结论
因素	318.98	3	106.33	3.76	显著
误差	395.46	14	28.25		
总和	714.44	17			

例9 对某地三所小学五年级男生身高进行了随机抽查，抽得身高(单位:cm)如表 9.19 所示. 问:

(1) 三所小学五年级男生身高有无显著差异?

(2) 三所小学五年级男生各自平均身高与方差的点估计数值($\alpha=0.05$)?

表 9.19

第一小学	128.1	134.1	133.1	138.9	140.8	127.4
第二小学	150.3	147.9	136.8	126.0	150.7	155.8
第三小学	140.6	143.1	144.5	143.7	148.5	146.4

解 (1) 以 μ_1,μ_2,μ_3 记三所小学五年级男生的平均身高，待检假设 $H_0:\mu_1=\mu_2=\mu_3$. 计算得

$$s=3,\quad n_1=n_2=n_3=6,\quad n=18.$$

$$S_T=\sum_{i=1}^{6}\sum_{j=1}^{3}x_{ij}^2-\frac{1}{18}T^2=1265.14,$$

$$S_A=\frac{1}{6}\sum_{j=1}^{3}x_{\cdot j}^2-\frac{1}{18}T^2=465.88,$$

$$S_E=S_T-S_A=799.26.$$

方差分析结果如表 9.20 所示. 因为 $F_{0.05}(2,15)=3.68<F_{比}=4.37$，故拒绝 H_0，认为三所小学五年级男生的平均身高有显著差异.

表 9.20

方差来源	平方和	自由度	均方	$F_{比}$	结论
因素	465.88	2	232.94	4.37	显著
误差	799.26	15	53.28		
总和	1265.14	17			

(2) $\hat{\mu}_1=\bar{x}_1=133.73, \hat{\mu}_2=\bar{x}_2=144.58, \hat{\mu}_3=\bar{x}_3=144.47$,

$$\hat{\sigma}^2=\frac{S_E}{n-s}=\frac{799.26}{15}=53.28.$$

例 10 为探讨教师对学生智力的估价是否影响学生智力发展的问题，任意地选取18 名学生进行试验，将这18 名学生随机地分为三组，每组6 名. 先测每名学生智商，然后对第一组学生宣称，他们的智力不大可能有较大提高；对第二组学生宣称，他们的智力会有中等程度的提高；对第三组学生宣称，他们的智力会有很大的提高. 一年后再对这些学生测试智商，得两次智商测试的差值如表9.21所示. 据此能否认为教师的评价会影响学生智力的发展($\alpha=0.05$)？

表 9.21

第一组	3	2	6	9	11	5
第二组	10	4	11	14	6	3
第三组	20	10	16	15	9	8

解 以 μ_1, μ_2, μ_3 记三组智商差值总体的均值，待检假设 $H_0: \mu_1=\mu_2=\mu_3$. 计算得

$$s=3, \quad n_1=n_2=n_3=5, \quad n=15,$$

$$S_T=\sum_{j=1}^{3}\sum_{i=1}^{6}x_{ij}^2-\frac{1}{18}T^2=422,$$

$$S_A=\frac{1}{6}\sum_{j=1}^{3}T_{\cdot j}^2-\frac{1}{18}T^2=156,$$

$$S_E=S_T-S_A=266.$$

方差分析结果如表 9.22 所示. 因为 $F_{0.05}(2,15)=3.68<F_{比}=4.40$，故拒绝 H_0，认为教师的估价显著影响学生智力的发展.

表 9.22

方差来源	平方和	自由度	均方	$F_{比}$	结论
因素	156	2	78	4.40	显著
误差	266	15	17.73		
总和	422	17			

例 11 有三台机床生产同一种产品，记录其 5 d 的产量如表 9.23所示. 问：在 $\alpha=0.05$ 下，机床对产量的影响是否显著？求出 $\hat{\sigma}^2,\mu_j,\delta_j\ (j=1,2,3)$及 $\mu_1-\mu_2,\mu_1-\mu_3,\mu_2-\mu_3$ 的置信度为0.95 的置信区间.

表 9.23

机床 1	48	45	56	51	48
机床 2	41	49	48	41	57
机床 3	65	54	72	51	64

解 以 μ_1,μ_2,μ_3 记各机床产量总体的均值，待检假设 $H_0:\mu_1=\mu_2=\mu_3$. 计算得

$$s=3,\quad n_1=n_2=n_3=5,\quad n=15,$$

$$S_T=\sum_{j=1}^{3}\sum_{i=1}^{5}x_{ij}^2-\frac{1}{15}T^2=42708-\frac{1}{15}\times 624100=1101.33,$$

$$S_A=\frac{1}{5}\sum_{j=1}^{3}T_{\cdot j}^2-\frac{1}{15}T^2=42167.2-\frac{1}{15}\times 624100=560.5,$$

$$S_E=S_T-S_A=540.83.$$

方差分析结果如表 9.24 所示. 因为 $F_{0.05}(2,12)=3.89<F_{比}=6.22$，故拒绝 H_0，认为机床对产量的影响是显著的.

表 9.24

方差来源	平方和	自由度	均方	$F_{比}$	结论
因素	560.5	2	280.25	6.22	显著
误差	540.83	12	45.07		
总和	1101.33	14			

作出如下估计：

$$\hat{\sigma}^2=\frac{S_E}{n-s}=\frac{540.83}{15-3}=45.07,\quad \hat{\mu}=\bar{x}=52.67,$$

$$\hat{\mu}_1=\bar{x}_1=49.6,\quad \hat{\mu}_2=\bar{x}_2=47.2,\quad \hat{\mu}_3=\bar{x}_3=61.2,$$

$$\hat{\delta}_1=\hat{\mu}_1-\hat{\mu}=-3.07,\quad \hat{\delta}_2=\hat{\mu}_2-\hat{\mu}=-5.47,$$
$$\hat{\delta}_3=\hat{\mu}_3-\hat{\mu}=8.53.$$
因为σ^2未知，由$\hat{\sigma}^2$代替，则均值差的置信度为$1-\alpha$的置信区间为
$$\left(\bar{x}_1-\bar{x}_2\mp t_{\alpha/2}(n-s)\sqrt{\hat{\sigma}^2}\Big/\sqrt{1/n_1+1/n_2}\right)=(\bar{x}_1-\bar{x}_2\mp 9.25),$$
所以，$\mu_1-\mu_2$的置信区间为
$$((49.6-47.2)\mp 9.25)=(-6.85,11.65),$$
$\mu_1-\mu_3$的置信区间为
$$((49.6-61.2)\mp 9.25)=(-20.85,-2.35),$$
$\mu_2-\mu_3$的置信区间为
$$((47.2-61.2)\mp 9.25)=(-23.25,-4.75).$$

例12 测量了四种不同类型外壳的彩色显像管的传导率，得传导率的观察值如表9.25所示.问:外壳类型对传导率有无显著影响($\alpha=0.05$)?

表9.25

类型1	143	141	150	146
类型2	152	144	137	143
类型3	134	136	133	129
类型4	129	128	134	129

解 以μ_1,μ_2,μ_3,μ_4记不同类型外壳总体的传导率均值.待检假设$H_0:\mu_1=\mu_2=\mu_3=\mu_4$.计算得
$$s=4,\quad n_1=n_2=n_3=n_4=4,\quad n=16,$$
$$S_T=\sum_{j=1}^{4}\sum_{i=1}^{4}x_{ij}^2-\frac{1}{16}T^2=305608-304704=904,$$
$$S_A=\frac{1}{4}\sum_{j=1}^{4}T_{\cdot j}^2-\frac{1}{16}T^2=305400-304704=696,$$
$$S_E=S_T-S_A=208.$$
方差分析结果如表9.26所示.因为$F_{0.05}(3,12)=3.49<F_{比}=$

13.38，故拒绝 H_0，认为外壳类型对传导率的影响是显著的.

表 9.26

方差来源	平方和	自由度	均方	$F_{比}$	结论
因素	696	3	232	13.38	显著
误差	208	12	17.33		
总和	904	15			

二、双因素方差分析

对双因素方差分析问题，要区别是无重复试验还是等重复试验. 无重复试验只需检验两个因素对试验结果有无显著影响，而等重复试验还要考察两个因素的交互作用对试验结果有无显著影响.

对双因素无重复试验的方差分析的计算，为了计算方便，S_T，S_A，S_B，S_E 也可以用以下公式得出：

$$S_T=\sum_{i=1}^{r}\sum_{j=1}^{s}x_{ij}^2-\frac{T^2}{rs},\quad S_A=\frac{1}{s}\sum_{i=1}^{r}T_{i\cdot}^2-\frac{T^2}{rs},$$

$$S_B=\frac{1}{r}\sum_{j=1}^{s}T_{\cdot j}^2-\frac{T^2}{rs},\quad S_E=S_T-S_A-S_B.$$

其中 $T=\sum_{i=1}^{r}\sum_{j=1}^{s}x_{ij}$，$T_{i\cdot}=\sum_{j=1}^{s}x_{ij}$，$T_{\cdot j}=\sum_{i=1}^{r}x_{ij}$.

对双因素等重复试验的方差分析，简化公式为

$$T=\sum_{i=1}^{r}\sum_{j=1}^{s}\sum_{k=1}^{t}x_{ijk},$$

$$T_{ij\cdot}=\sum_{k=1}^{t}x_{ijk},\quad i=1,2,\cdots,r \text{ 且 } j=1,2,\cdots,s,$$

$$T_{i\cdot\cdot}=\sum_{j=1}^{s}\sum_{k=1}^{t}x_{ijk},\quad i=1,2,\cdots,r,$$

$$T_{\cdot j\cdot}=\sum_{i=1}^{r}\sum_{k=1}^{t}x_{ijk},\quad j=1,2,\cdots,s,$$

则 S_T，S_A，S_B，S_{AB}，S_E 可以由以下公式得出：

$$S_T=\sum_{i=1}^{r}\sum_{j=1}^{s}\sum_{k=1}^{t}x_{ijk}^2-\frac{T^2}{rst},$$

$$S_A=\frac{1}{st}\sum_{i=1}^{r}T_{i\cdot\cdot}^2-\frac{T^2}{rst},\quad S_B=\frac{1}{rt}\sum_{j=1}^{s}T_{\cdot j\cdot}^2-\frac{T^2}{rst},$$

$$S_{AB}=\left(\frac{1}{t}\sum_{i=1}^{r}\sum_{j=1}^{s}T_{ij\cdot}^2-\frac{T^2}{rst}\right)-S_A-S_B,$$

$$S_E=S_T-S_A-S_B-S_{AB}.$$

例13 某工厂在生产一种产品时使用了三种不同的催化剂和四种不同的原料，每种搭配都做一次试验，测得产品的抗压强度(单位：MPa)数据如表9.27所示. 试在$\alpha=0.05$下检验不同催化剂和原料对抗压强度有无显著影响.

表9.27

		催化剂		
		A_1	A_2	A_3
原料	B_1	31	33	35
	B_2	34	36	37
	B_3	35	37	39
	B_4	39	38	42

解 设α_i为因素A在水平A_i的效应，β_j为因素B在水平β_j的效应. 待检假设

$$H_{01}:\alpha_1=\alpha_2=\alpha_3=0,\quad H_{02}:\beta_1=\beta_2=\beta_3=\beta_4=0.$$

因为$r=3,s=4$，所以

$$S_T=15940-\frac{1}{3\times4}\times436^2=98.67,$$

$$S_A=\frac{1}{4}\times63466-\frac{1}{3\times4}\times436^2=25.17,$$

$$S_B=\frac{1}{3}\times47732-\frac{1}{3\times4}\times436^2=69.34,$$

$$S_E=S_T-S_A-S_B=4.16.$$

方差分析结果如表9.28所示. 因为

$$F_{0.05}(2,6)=5.14<F_{比}=18.16,$$

$$F_{0.05}(3.6)=4.76<F_{比}=33.35,$$

所以拒绝 H_{01} 和 H_{02}，认为催化剂和原料的影响都是显著的.

表 9.28

方差来源	平方和	自由度	均方	$F_{比}$	结论
因素 A	25.17	2	12.585	18.16	显著
因素 B	69.34	3	23.113	33.35	显著
误 差	4.16	6	0.693		
总 和	98.67	11			

例 14 在 B_1,B_2,B_3,B_4 四台纺织机器中，用三种不同的加压水平 A_1,A_2,A_3，在每种加压水平和每台机器中各取一个试样测量，得纱支强度(单位：Pa)如表 9.29 所示. 问：不同加压水平及不同机器之间纱支强度有无显著差异($\alpha=0.05$)？

表 9.29

		机器			
		B_1	B_2	B_3	B_4
加压水平	A_1	1577	1690	1800	1642
	A_2	1535	1640	1783	1621
	A_3	1592	1652	1810	1663

解 以 α_i $(i=1,2,3)$ 记因素 A 在水平 A_i 的效应，β_j 为因素 B 在水平 β_j 的效应. 待检假设

$$H_{01}:\alpha_1=\alpha_2=\alpha_3=0,\quad H_{02}:\beta_1=\beta_2=\beta_3=\beta_4=0.$$

因为 $r=3,s=4$，所以

$$S_T=\sum_{i=1}^{3}\sum_{j=1}^{4}x_{ij}^2-\frac{1}{3\times 4}T^2=88650-\frac{1}{12}\times 20005=86982.92,$$

$$S_A=\frac{1}{4}\sum_{i=1}^{3}T_{i\cdot}^2-\frac{1}{12}T^2=3000.67,$$

$$S_B=\frac{1}{3}\sum_{j=1}^{4}T_{\cdot j}^2-\frac{1}{12}T^2=82619.58,$$

$$S_E=S_T-S_A-S_B=1362.67.$$

方差分析结果如表9.30所示.因为

$$F_{0.05}(2,6)=3.14<F_{比}=6.61,$$

$$F_{0.05}(3,6)=4.76<F_{比}=121.26,$$

所以拒绝H_{01}和H_{02},认为不同加压水平对纱支强度影响显著,不同机器对纱支强度的影响高度显著.

表9.30

方差来源	平方和	自由度	均 方	$F_{比}$	结 论
因素A	3000.67	2	1500.33	6.61	显著
因素B	82619.58	3	27539.86	121.26	高度显著
误 差	1362.67	6	227.11		
总 和	86982.92	11			

例15 设有6种不同品种的种子和5种不同的施肥方案,在30块相同面积的地块上,分别对种子与肥料的不同搭配进行试验,收获量(单位:kg)如表9.31所示.试在$\alpha=0.05$下检验种子的品种与施肥方案的不同对收获量是否有显著影响.

表9.31

		品 种					
		1	2	3	4	5	6
施肥方案	1	12.0	11.5	11.5	11.0	9.5	9.3
	2	10.8	11.4	12.0	11.1	9.6	9.7
	3	13.2	13.1	12.5	11.4	12.4	10.4
	4	14.0	14.0	14.0	12.3	11.5	9.5
	5	11.6	13.0	14.2	14.3	13.7	12.0

解 以α_i记种子因素A的不同品种的效应,以β_j记施肥因素B的不同方案的效应,待检假设

$H_{01}:\alpha_1=\alpha_2=\alpha_3=\alpha_4=\alpha_5=\alpha_6$, $H_{02}:\beta_1=\beta_2=\beta_3=\beta_4=\beta_5$.

因为$r=6,s=5,T=356.5,T_{1\cdot}=61.6,T_{2\cdot}=63,T_{3\cdot}=64.2,T_{4\cdot}=60.1,T_{5\cdot}=56.7,T_{6\cdot}=50.9,T_{\cdot 1}=64.8,T_{\cdot 2}=64.6,T_{\cdot 3}=73,T_{\cdot 4}=75.3,T_{\cdot 5}=78.8$,所以

$$S_T=\sum_{i=1}^{6}\sum_{j=1}^{5}x_{ij}^2-\frac{1}{5\times 6}T^2=4303.45-4236.4=67.05,$$

$$S_A=\frac{1}{5}\sum_{i=1}^{6}T_{i\cdot}^2-\frac{1}{5\times 6}T^2=4260.58-4236.4=24.18,$$

$$S_B=\frac{1}{6}\sum_{j=1}^{5}T_{\cdot j}^2-\frac{1}{5\times 6}T^2=4263.46-4236.4=27.06,$$

$$S_E=S_T-S_A-S_B=67.05-24.18-27.06=15.81.$$

方差分析结果如表 9.32 所示. 因为

$$F_{0.05}(5,20)=4.10<F_{比}=6.37,$$

$$F_{0.05}(4,20)=4.43<F_{比}=8.9,$$

故拒绝 H_{01} 和 H_{02}，认为种子的品种与施肥方案对收获量的影响都是显著的.

表 9.32

方差来源	平方和	自由度	均方	$F_{比}$	结论
因素 A	24.18	5	4.836	6.37	高度显著
因素 B	27.06	4	6.765	8.9	高度显著
误 差	15.81	20	0.791		
平方和	67.05	29			

例 16 为考察某种合金中碳的质量分数(因素 A)与锑-铝质量分数之和(因素 B)对合金强度(单位:MPa)的影响，对因素 A 取三个水平，因素 B 取四个水平，在每对组合下作一次试验，得强度数据如表 9.33 所示. 试在 $\alpha=0.01$ 下检验因素 A 和 B 的效应是否显著.

表 9.33

A \ B	3.3%	3.4%	3.5%	3.6%
0.03%	63.1	63.9	65.6	66.8
0.04%	65.1	66.4	67.8	69.0
0.05%	67.2	71.0	71.9	73.5

解 以α_i记因素A的水平A_i的效应，β_i记因素B_i的效应. 待检假设

$$H_{01}:\alpha_1=\alpha_2=\alpha_3=0,\quad H_{02}:\beta_1=\beta_2=\beta_3=0.$$

因为$r=3, s=4, T=811.3, T_{1\cdot}=259.4, T_{2\cdot}=268.3, T_{3\cdot}=283.6, T_{\cdot 1}=195.4, T_{\cdot 2}=201.3, T_{\cdot 3}=205.3, T_{\cdot 4}=209.3$，所以

$$S_T=\sum_{i=1}^{3}\sum_{j=1}^{4}x_{ij}^2-\frac{1}{rs}T^2=54963.93-\frac{1}{12}\times 811.3^2=113.29,$$

$$S_A=\frac{1}{4}\sum_{i=1}^{3}x_{i\cdot}^2-\frac{1}{rs}T^2=54925.55-54850.64=74.91,$$

$$S_B=\frac{1}{3}\sum_{j=1}^{4}x_{\cdot j}^2-\frac{1}{rs}T^2=54885.81-54850.64=35.17,$$

$$S_E=S_T-S_A-S_B=3.21.$$

方差分析结果如表9.34所示. 因为

$$F_{0.01}(2,6)=10.9<F_{比}=70.02,$$

$$F_{0.01}(3,6)=9.78<F_{比}=21.91,$$

所以拒绝H_{01}和H_{02}，认为合金中碳含量和锑-铝含量对合金强度的影响都是显著的.

表9.34

方差来源	平方和	自由度	均方	$F_{比}$	结论
因素A	74.91	2	37.455	70.02	显著
因素B	35.17	3	11.723	21.91	显著
误 差	3.21	6	0.535		
总 和	113.29	11			

例17 某细纱车间记录了甲、乙、丙三名工人在四组机台上操作的产量，其数据如表9.35所示. 试在$\alpha=0.05$下检验：不同工人操作之间差异是否显著，机床之间差异是否显著，两个因素的交互作用是否显著.

解 这是双因素等重复方差分析问题. 待检假设

$$H_{01}:\alpha_1=\alpha_2=\alpha_3=0,\quad H_{02}:\beta_1=\beta_2=\beta_3=\beta_4=0,$$

$$H_{03}:\gamma_{11}=\gamma_{12}=\cdots=\gamma_{34}=0.$$

表 9.35

		工人								
		甲(A_1)			乙(A_2)			丙(A_3)		
机台	B_1	15	15	17	19	19	16	16	18	21
	B_2	17	17	17	15	15	15	19	22	22
	B_3	15	17	16	18	17	16	18	18	18
	B_4	18	20	22	15	16	17	17	17	17

因为 $r=3,\quad s=4,\quad t=3,\quad T=627,$

$$\sum_{i=1}^{r}\sum_{j=1}^{s}\sum_{k=1}^{t}x_{ijk}^2=11065,$$

所以
$$S_T=\sum_{i=1}^{3}\sum_{j=1}^{4}\sum_{k=1}^{3}x_{ijk}^2-\frac{1}{rst}T^2=11065-\frac{1}{36}\times 627^2$$
$$=144.75,$$

$$S_A=\frac{1}{st}\sum_{i=1}^{r}T_{i..}^2-\frac{1}{rst}T^2=\frac{1}{12}\times 131369-10920.25$$
$$=27.17,$$

$$S_B=\frac{1}{rt}\sum_{j=1}^{s}T_{.j.}^2-\frac{1}{rst}T^2=\frac{1}{9}\times 98307-10920.25=2.75,$$

$$S_{AB}=\left(\frac{1}{t}\sum_{i=1}^{r}\sum_{j=1}^{s}T_{ij.}^2-\frac{1}{rst}\right)-S_A-S_B=73.5,$$

$$S_E=S_T-S_A-S_B-S_{AB}=41.33.$$

方差分析结果如表 9.36 所示. 因为 $F_{0.05}(2,24)=3.40<F_{比}=7.89$,所以工人操作因素的影响显著;又因为 $F_{0.05}(3,24)=3.01>F_{比}=0.53$,所以机床因素的影响不显著;又因为 $F_{0.05}(6,24)=2.51<F_{比}=7.11$,所以两因素交互作用影响显著.

表 9.36

方差来源	平方和	自由度	均方	$F_{比}$	结论
因素 A	27.17	2	13.58	7.89	显著
因素 B	2.75	3	0.92	0.53	不显著
因素 AB	73.5	6	12.25	7.11	显著
误 差	41.33	24	1.7		
总 和	144.75	35			

例18 为了考察对纤维弹性测量的误差，今对同一批原料，由四个检测站（A_1, A_2, A_3, A_4）同时测量，每站各出一名检验员（B_1, B_2, B_3, B_4），轮流使用各站设备作重复测量，得数据如表 9.37 所示. 试在 $\alpha=0.05$ 下检验：不同检测站，不同检验员以及他们的交互作用对误差的影响是否显著.

表 9.37

	B_1	B_2	B_3	B_4	行和
A_1	71,73	72,73	75,73	77,75	589
A_2	73,75	76,74	78,77	76,74	603
A_3	76,73	79,77	74,75	74,73	601
A_4	75,73	73,72	70,71	69,69	572
列和	589	596	593	587	2365
$\sum_{i=1}^{4}\sum_{k=1}^{2} x_{ijk}$	43383	44448	44009	43133	
$\sum_{i=1}^{4}\left(\sum_{k=1}^{2} x_{ijk}\right)^2$	86745	88886	88011	86253	

解 待检假设

$$H_{01}:\alpha_1=\alpha_2=\alpha_3=\alpha_4=0, \quad H_{02}:\beta_1=\beta_2=\beta_3=\beta_4=0,$$

$$H_{03}:\gamma_{11}=\gamma_{12}=\cdots=\gamma_{44}=0.$$

因为 $$r=4, \quad s=4, \quad t=2,$$

所以 $$S_T=184.719, \quad S_A=76.095, \quad S_B=6.095,$$

$$S_{AB}=79.03, \quad S_E=23.50.$$

方差分析结果如表 9.38 所示. 因为 $F_{0.05}(3,16)=3.24<F_{比}=17.27$,所以不同检测站的影响高度显著;因为 $F_{0.05}(3,16)=3.24>F_{比}=1.39$,所以不同检验员的影响不显著;因为 $F_{0.05}(9,16)=2.54<F_{比}=5.98$,所以两因素的交互作用影响显著.

表 9.38

方差来源	平方和	自由度	均方	$F_{比}$	结论
因素 A	76.095	3	25.365	17.27	显著
因素 B	6.095	3	2.032	1.39	不显著
因素 AB	79.03	9	8.78	5.98	显著
误 差	23.50	16	1.47		
总 和	184.719	31			

例 19 发电机的寿命与制造材料及使用地点的温度有关. 今选取三种不同的材料及两种不同的温度作重复试验,得数据如表 9.39 所示. 试在 $\alpha=0.05$ 下,检验不同材料,不同温度及交互作用对寿命的影响是否显著.

表 9.39

		温度					
		B_1(10 ℃)			B_2(18 ℃)		
材料	A_1	136	150	176	50	54	64
	A_2	150	162	171	76	88	91
	A_3	138	109	140	68	62	77

解 是双因素等重复试验的方差分析. 待检假设

$$H_{01}:\alpha_1=\alpha_2=\alpha_3=0,\quad H_{02}:\beta_1=\beta_2=0,$$

$$H_{03}:\gamma_{11}=\gamma_{12}=\cdots=\gamma_{32}=0.$$

因为 $$r=3,\quad s=2,\quad t=3,$$

$$T_{1\cdot\cdot}=630,\quad T_{2\cdot\cdot}=738,\quad T_{3\cdot\cdot}=594,$$

$$T_{\cdot1\cdot}=1332,\quad T_{\cdot2\cdot}=630,\quad T=1962,$$

$$\sum_{i=1}^{3}\sum_{j=1}^{2}\sum_{k=1}^{3}x_{ijk}^2=246192.$$

$$S_T=246192-213858=32334,$$

$$S_A=215730-213858=1872,$$

$$S_B=241236-213858=23378,$$

所以 $$S_{AB}=244200-21385-1872-23378=5092,$$

$$S_E=S_T-S_A-S_B-S_{AB}=1992,$$

方差分析结果如表 9.40 所示. 因为 $F_{0.05}(2,12)=3.89<F_{比}=5.64$,所以材料对寿命的影响显著;因为 $F_{0.05}(1,12)=4.75<F_{比}=140.85$,所以温度对寿命的影响高度显著;因为 $F_{0.05}(2,12)=3.89<F_{比}=15.34$,所以交互作用对寿命的影响显著.

表 9.40

方差来源	平方和	自由度	均方	$F_{比}$	结论
因素 A	1872	2	936	5.64	显著
因素 B	23378	1	23378	140.85	显著
因素 AB	5092	2	2546	15.34	显著
误差	1992	12	166		
总和	32334	17			

例20 某职校在招收在职生时,为考察年龄与工龄对成绩(百分制)的影响,各取两个水平进行重复试验,得数据如表 9.41 所示. 试用方差分析法确定,招收在职生的最佳年龄与工龄.

表 9.41

	B_1(5 年以下)					B_2(5 年以上)				
A_1(25 岁以下)	86	87	76	79	85	82	93	82	88	91
A_2(25 岁以上)	77	82	84	90	76	82	82	80	75	79

解 这是双因素等重复试验的方差分析. 待检假设

$$H_{01}:\alpha_1=\alpha_2=0,\quad H_{02}:\beta_1=\beta_2=0,$$

$$H_{03}:\gamma_{11}=\gamma_{12}=\gamma_{21}=\gamma_{22}=0.$$

因为 $r=2,\quad s=2,\quad t=5,$

$$T_{11\cdot}=413,\quad T_{21\cdot}=409,\quad T_{12\cdot}=436,\quad T_{22\cdot}=398,$$

$$T_{1\cdot\cdot}=849,\quad T_{2\cdot\cdot}=807,\quad T_{\cdot 1\cdot}=822,\quad T=1656,$$

$$\sum T_{i\cdot\cdot}^2=1372050,\ \sum T_{\cdot j\cdot}^2=1371240,\ \sum_{i=1}^{2}\sum_{j=1}^{2}T_{ij\cdot}^2=686350,$$

所以

$$S_T=\sum_{i=1}^{2}\sum_{j=1}^{2}\sum_{k=1}^{5}x_{ijk}^2-\frac{1}{rst}T^2=137630-137116.8=513.2,$$

$$S_A=\frac{1}{st}\sum_{i=1}^{r}T_{i\cdot\cdot}^2-\frac{1}{rst}T^2=137205-137116.8=88.2,$$

$$S_B=\frac{1}{rt}\sum_{j=1}^{s}T_{\cdot j\cdot}^2-\frac{1}{rst}T^2=137124-137116.8=7.2,$$

$$S_{AB}=\left(\frac{1}{t}\sum_{i=1}^{r}\sum_{j=1}^{s}T_{ij\cdot}^2-\frac{1}{rst}T^2\right)-S_A-S_B$$
$$=153.2-95.4=57.8.$$

$$S_E=S_T-S_A-S_B-S_{AB}=360.$$

方差分析结果如表 9.42 所示. 因为

$$F_{0.05}(1,16)=4.49>F_{比}=3.94,$$
$$F_{0.05}(1,16)=4.49>F_{比}=0.32,$$
$$F_{0.05}(1.16)=4.49>F_{比}=2.58,$$

所以年龄、工龄以及交互作用对学习成绩都无显著影响. 但从平均成绩来看,25 岁以下者平均成绩为 84.9 分,25 岁以上者平均成绩为 80.7 分,所以,以招收 25 岁以下者较优.

表 9.42

方差来源	平方和	自由度	均方	$F_{比}$	结论
因素 A	88.2	1	88.2	3.94	不显著
因素 B	7.2	1	7.2	0.32	不显著
因素 AB	57.8	1	57.8	2.58	不显著
误差	360	16	22.5		
总和	513.2	19			

第二节 回归分析

主要内容

回归分析是处理多个变量之间相互关系的一种数学方法，可分为一元线性回归与多元线性回归以及可化为线性回归的一元非线性回归.回归分析从数据出发，提供建立变量之间相关关系的表达式——经验公式，给出相关性检验规则，并运用经验公式达到预测与控制目的.

一、一元线性回归

对于一个问题中的两个变量x与y(其中x是一个普通变量，y是一个随机变量)，若对于x的每一个确定值，y有它的分布，则称x与y存在相关关系.当x取定时，y的数学期望$E(y)$是x的函数，称为y关于x的回归.

设$y \sim N(\mu(x), \sigma^2)$，$\mu(x)=E(y|x=x)=\hat{y}$是$y$的估计值，又称$y$关于$x$的回归方程.

若$\mu(x)=a+bx$，即$y \sim N(a+bx, \sigma^2)$，则估计$\mu(x)$的问题称为一元线性回归问题.由样本估计$\hat{a}$和$\hat{b}$，得到方程$\hat{y}=\hat{a}+\hat{b}x$，称之为y关于x的一元线性回归方程，其图形称为回归直线.

1. a，b的最小二乘估计

对一组样本值(x_1, y_1)，$(x_2, y_2,)$，…，(x_n, y_n)，画出散点图.应适当选取a，b作直线$\hat{y}=a+bx$，使直线与回归直线最接近，这时离差平方和$\sum_{i=1}^{n}[y_i-(a+bx_i)]^2$达到最小.

对函数 $$Q(a,b)=\sum_{i=1}^{n}[y_i-(a+bx_i)]^2$$

取关于 a 和 b 的偏导数，令其等于零，得方程组

$$\begin{cases} na+\left(\sum_{i=1}^{n}x_i\right)b=\sum_{i=1}^{n}y_i, \\ \left(\sum_{i=1}^{n}x_i\right)a+\left(\sum_{i=1}^{n}x_i^2\right)b=\sum_{i=1}^{n}x_iy_i, \end{cases}$$

解得 $\begin{cases} \hat{b}=\dfrac{l_{xy}}{l_{xx}}, \\ \hat{a}=\bar{y}-\hat{b}\bar{x}, \end{cases}$ 则 $\hat{y}=\hat{a}+\hat{b}x$ 为所求线性回归方程，其中

$$l_{xx}=\sum_{i=1}^{n}(x_i-\bar{x})^2=\sum_{i=1}^{n}x_i^2-n\bar{x}^2,$$

$$l_{xy}=\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})=\sum_{i=1}^{n}x_iy_i-n\bar{x}\bar{y}.$$

2. 方差 σ^2 的估计

方差 $\sigma^2=E(y-a-bx)^2$ 的无偏估计量是

$$\hat{\sigma}^2=\frac{1}{n-2}(l_{yy}-\hat{b}l_{xy}),$$

其中 $$l_{yy}=\sum_{i=1}^{n}(y_i-\bar{y})^2=\sum_{i=1}^{n}y_i^2-n\bar{y}^2.$$

3. 线性回归的分析

(1) 线性假设的显著性检验　待检假设 $H_0:b=0$. 常用 t 检验法、F 检验法、相关系数检验法.

t 检验法　$(\hat{b}-b)\sqrt{l_{xx}}\big/\hat{\sigma}\sim t(n-2)$，当 H_0 为真时，$b=0$，有统计量

$$T=\frac{\hat{b}}{\hat{\sigma}}\sqrt{l_{xx}}\sim t(n-2),$$

且 $E(\hat{b})=b=0$，则 H_0 的拒绝域为

$$|t|=\frac{|\hat{b}|}{\hat{\sigma}}\sqrt{l_{xx}}\geqslant t_{\alpha/2}(n-2).$$

F 检验法　当 H_0 为真时，统计量

$$F=\frac{U}{Q/(n-2)}\sim F(1,n-2),$$

则 H_0 的拒绝域为

$$F\geqslant F_{\alpha}(1,n-2).$$

当 H_0 被拒绝时，认为回归效果是显著的(式中，$U=l_{xy}^2/l_{xx}$，$Q=l_{yy}-U$)．

(2) 系数 b 的置信区间　当回归效果显著时，b 的置信度 $1-\alpha$ 的置信区间为

$$\left(\hat{b}-t_{\alpha/2}(n-2)\cdot\hat{\sigma}\Big/\sqrt{l_{xx}},\hat{b}+t_{\alpha/2}(n-2)\cdot\hat{\sigma}\Big/\sqrt{l_{xx}}\right).$$

(3) 预测　预测就是当给定 x_0 时，对 y 的取值作点估计或区间估计．

当给定 x_0 时，y_0 的预测值为 $\hat{y}_0=\hat{a}+\hat{b}x_0$．

y_0 的置信度为 $1-\alpha$ 的置信区间为

$$\left(y_0\mp t_{\alpha/2}(n-3)\hat{\sigma}\sqrt{1+1/n+(x_0-\bar{x})^2/l_{xx}}\right).$$

当 x_0 与 $\bar{x}$ 越接近，在给定的样本观察值与置信度下，预测区间宽度越窄，预测也越精确(见图 9.2)．

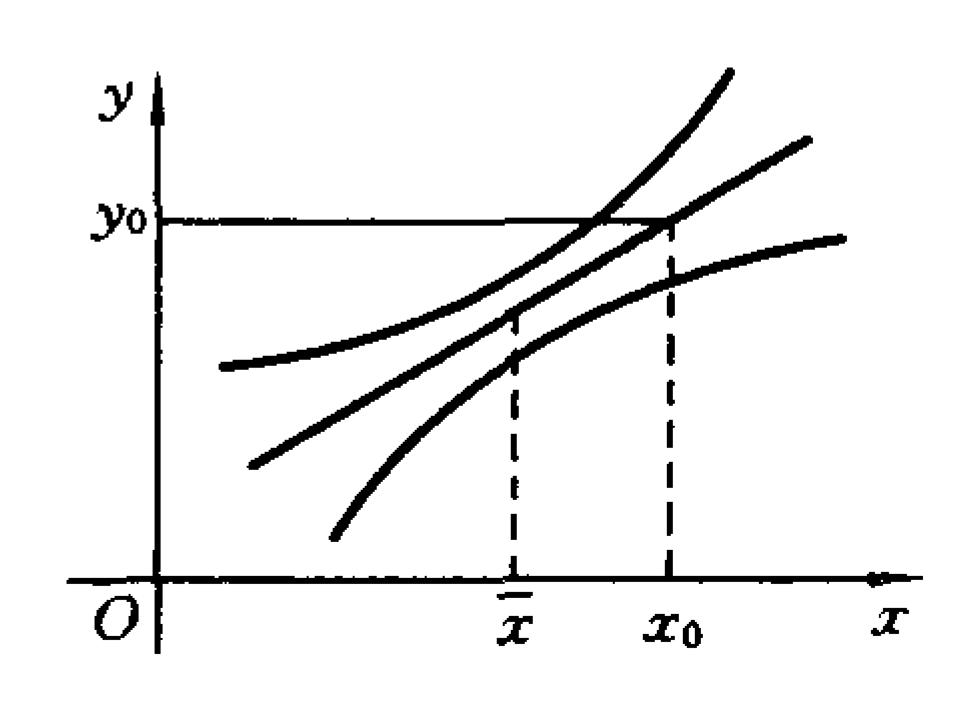
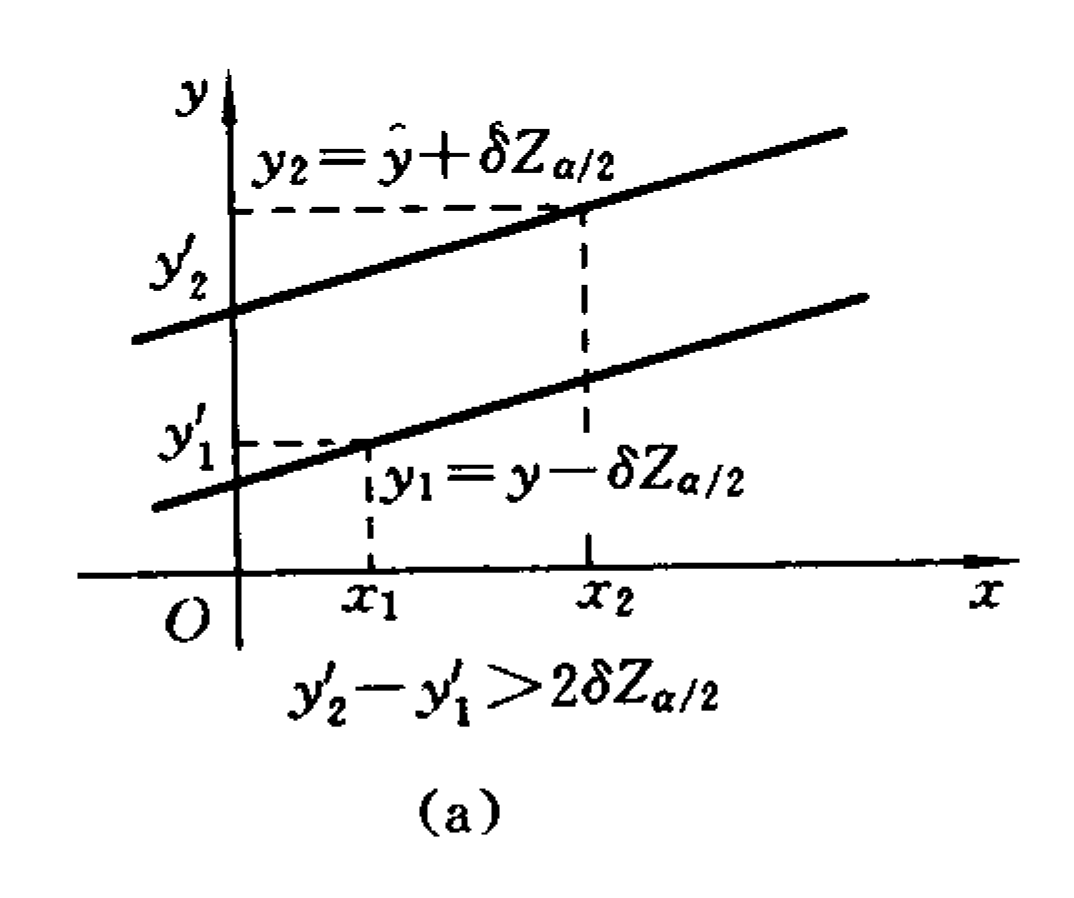


图 9.2

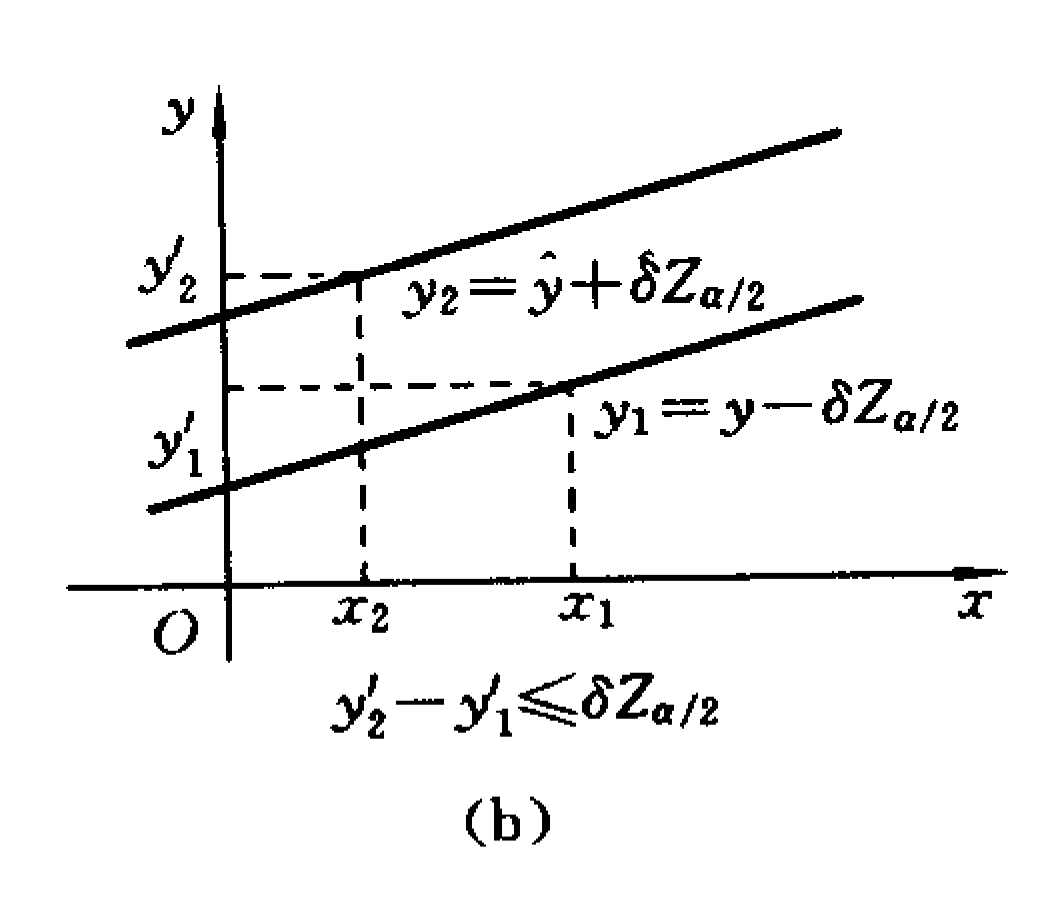
(4) 控制　如果给定了置信度 $1-\alpha$，要求出 x_1，x_2，使当 $x_1<x<x_2$ 时，x 所对应的观察值落在事先确定的区间 (y_1',y_2') 内的概率不小于 $1-\alpha$. 当 n 很大时，令

$$y_1'=\hat{y}-\hat{\sigma}Z_{\alpha/2}=\hat{a}+\hat{b}x-\hat{\sigma}Z_{\alpha/2},$$
$$y_2'=\hat{y}+\hat{\sigma}Z_{\alpha/2}=\hat{a}+\hat{b}x+\hat{\sigma}Z_{\alpha/2},$$

即可解出控制 x 的上、下限．但必须注意，(y_1',y_2') 的长度一定要大于 $2\hat{\sigma}Z_{\alpha/2}$(见图 9.3(a))．



(a)



(b)

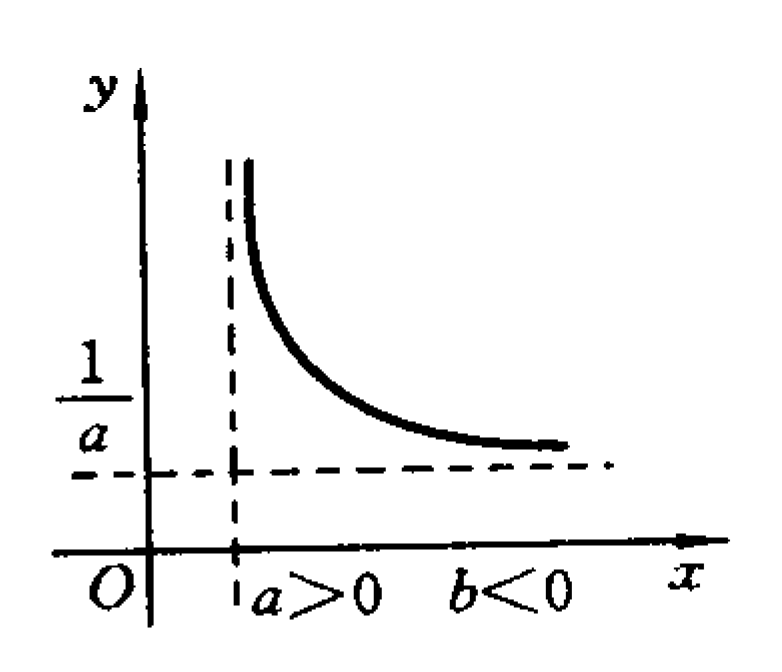
图 9.3

二、可化为线性回归的一元非线性回归

常见的可化为线性回归的非线性回归有以下几种.

1. 双曲线　$1/y=a+b/x$ 型(见图 9.4)

令 $u=1/x$，$v=1/y$，方程化为 $v=a+bu$. 由 (x_i,y_i) 算出 (u_i,v_i)，$i=1,2,\cdots,n$，估计参数 $\hat{a}$，$\hat{b}$，故有 $1/y=\hat{a}+\hat{b}/x$.



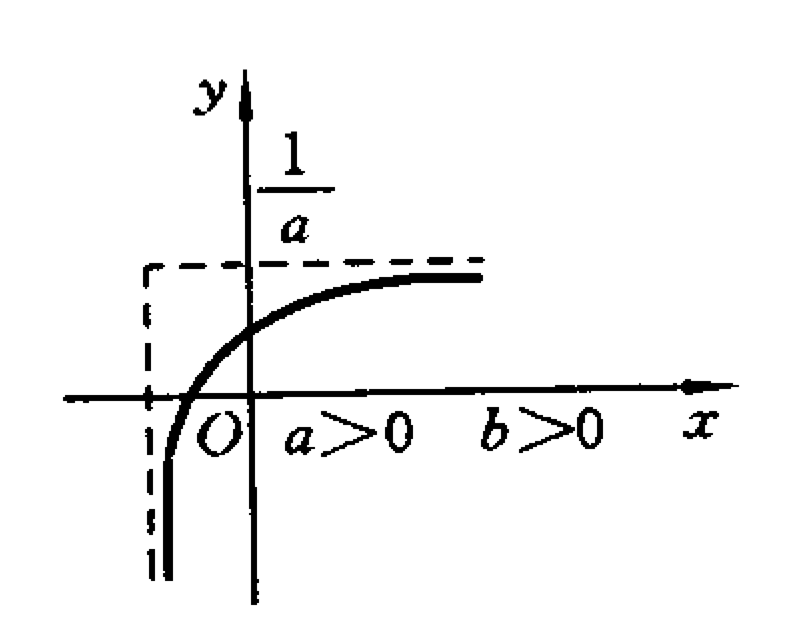
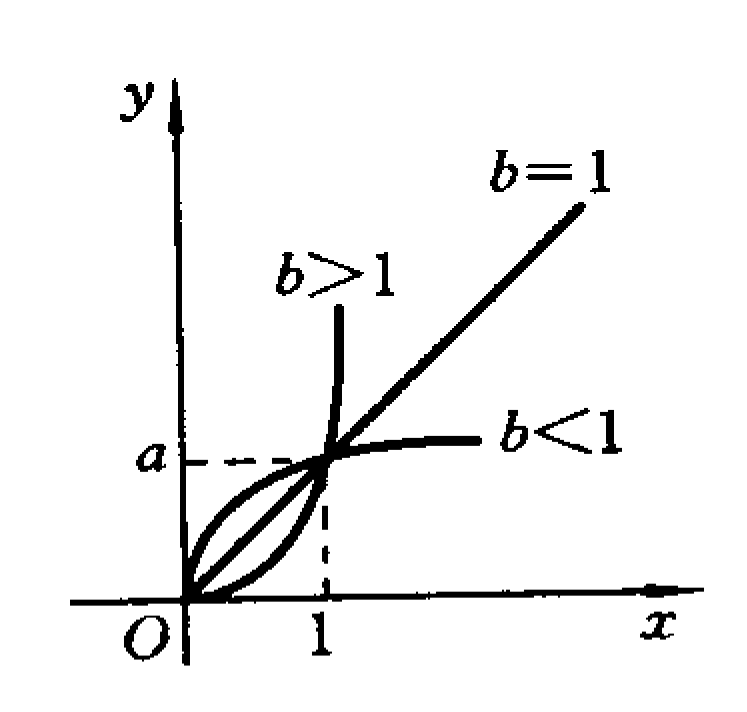


图 9.4

2. 幂函数曲线　$y=ax^b$ $(x>0,b>0)$型(见图 9.5)



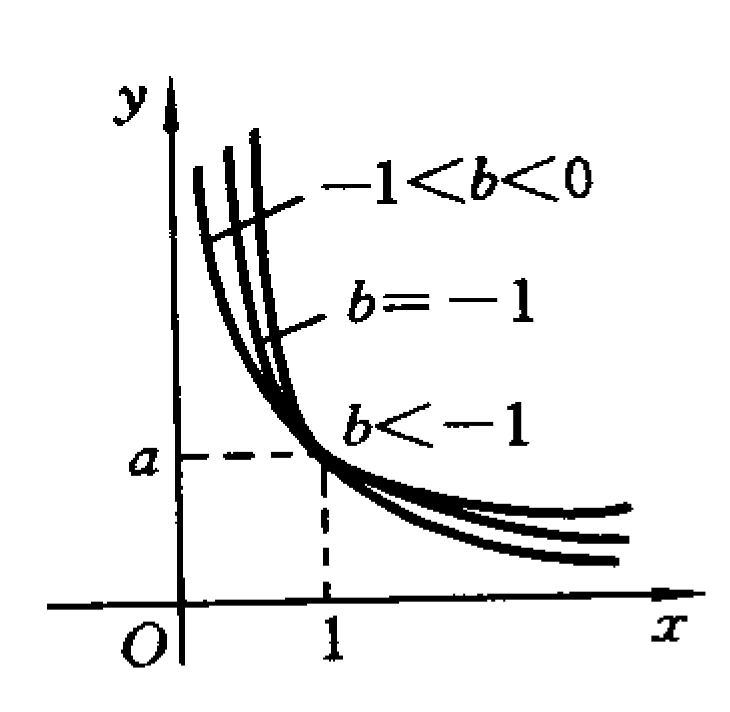


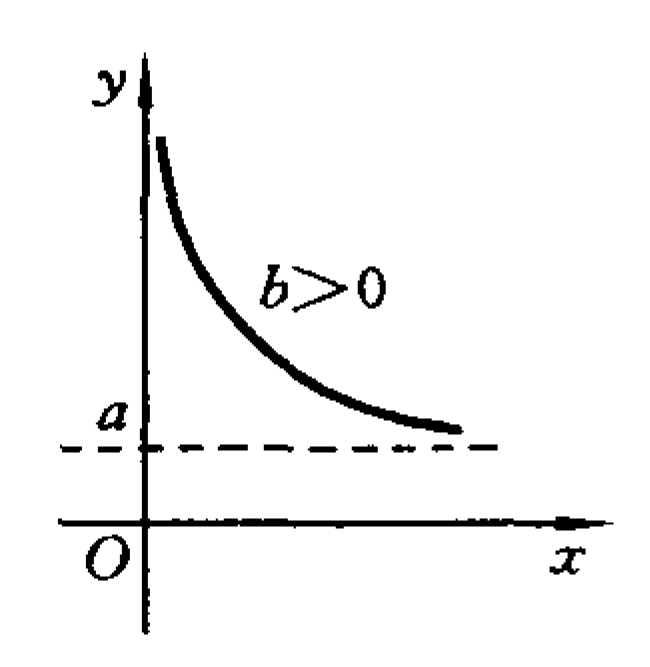
图 9.5

取对数，得 $\ln y=\ln a+b\ln x$，

令 $u=\ln x, v=\ln y, A=\ln a$，则有直线方程 $v=A+bu$.

由 (x_i, y_i) 算出 (u_i, v_i)，求出 $\hat{A}, \hat{b}$，再由 $\hat{a}=e^{\hat{A}}$，于是 $y=\hat{a}x^{\hat{b}}$.

3. 倒指数曲线　$y=ae^{b/x}$ $(a>0)$ 型（见图 9.6）



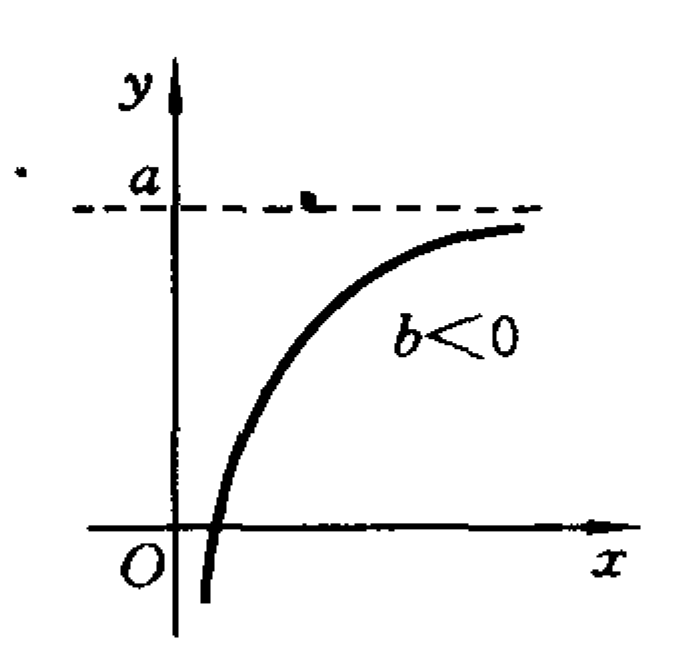


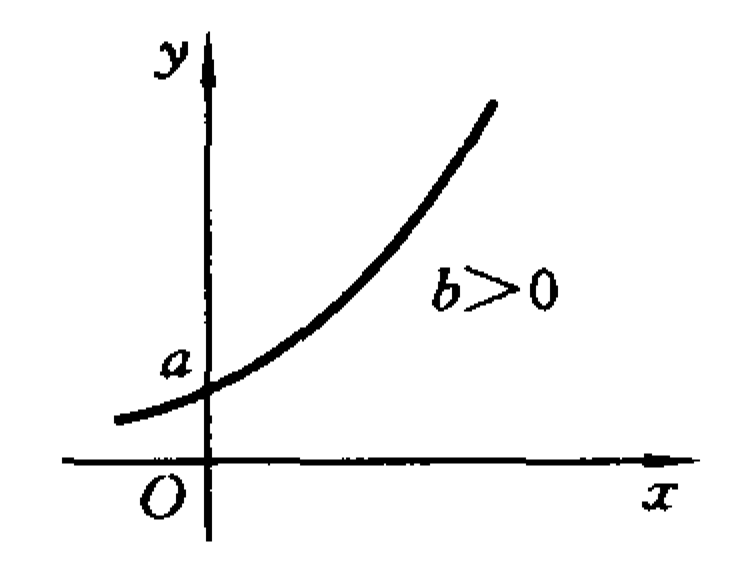
图 9.6

取对数，得 $\ln y=\ln a+b/x$，

令 $u=1/x, v=\ln y, A=\ln a$，化为直线方程 $v=A+bu$.

按幂函数曲线的方式计算得 $\hat{A}, \hat{b}, \hat{a}=e^{\hat{A}}$，于是得出 $y=\hat{a}e^{\hat{b}/x}$.

4. 指数曲线　$y=ae^{bx}$ $(a>0)$ 型（见图 9.7）



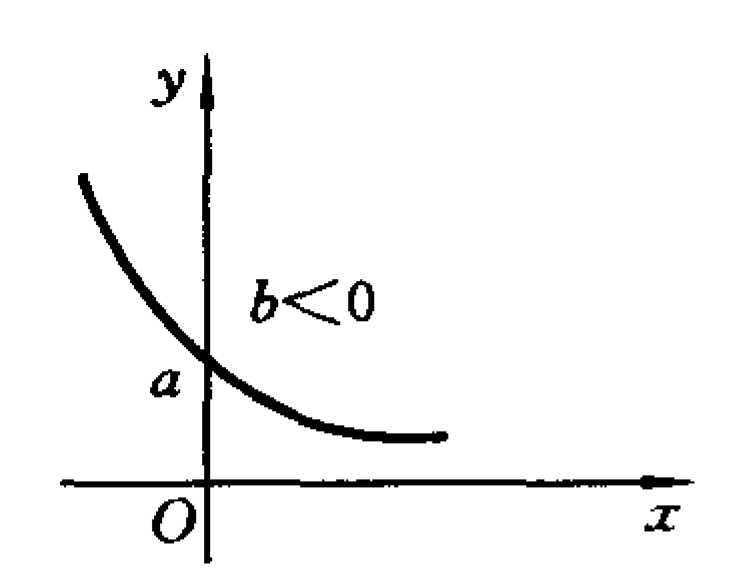


图 9.7

取对数，得 $\ln y=\ln a+bx$，

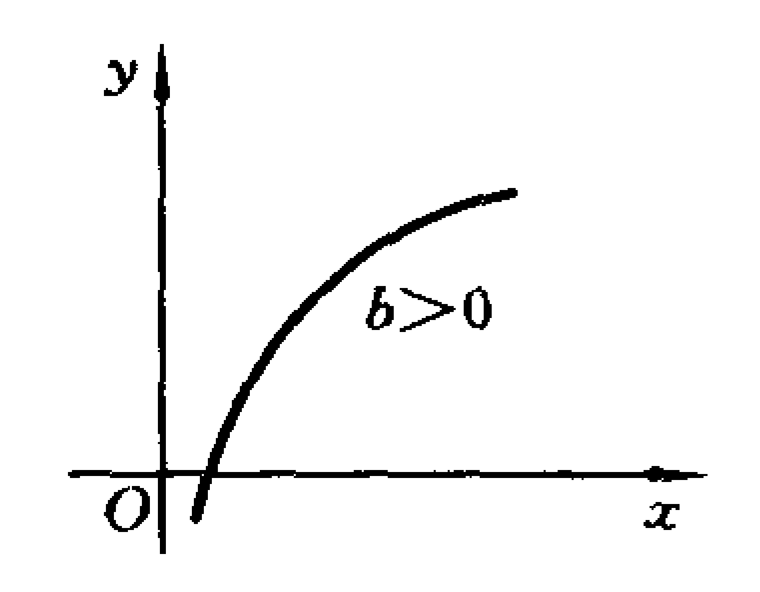
令 $v=\ln y, A=\ln a$，用前述方法求出直线方程 $v=A+bx$ 的估计值 $\hat{A}$ 和 $\hat{b}$，由 $\hat{a}=e^{\hat{A}}$，即得 $y=\hat{a}e^{\hat{b}x}$.

5. 对数函数　$y=a+b\lg x$ $(x>0)$ 型（见图 9.8）

令 $u=\lg x$，则有 $y=a+bu$，由数据算得 $\hat{a}, \hat{b}$，即有 $y=\hat{a}+\hat{b}\lg x$.

6. S 形曲线　$y=\dfrac{1}{a+be^{-x}}$ 型（见图9.9）

令 $u=e^{-x}, y=v$，则有直线方程



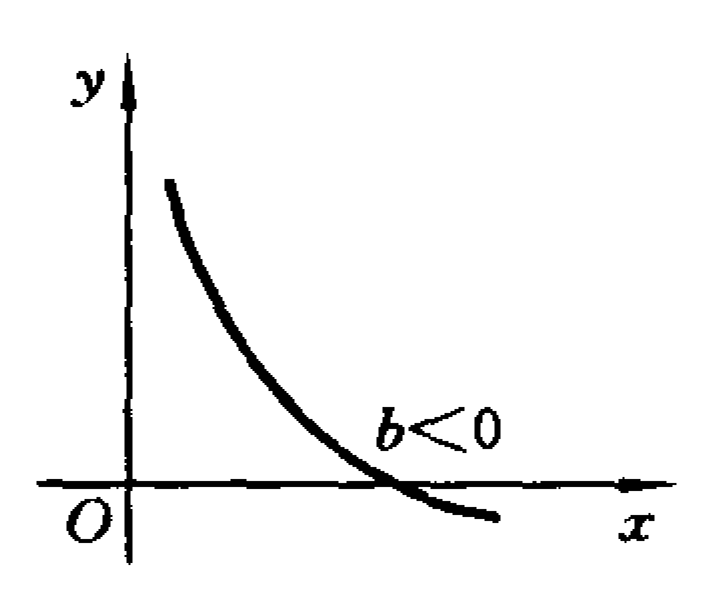


图 9.8

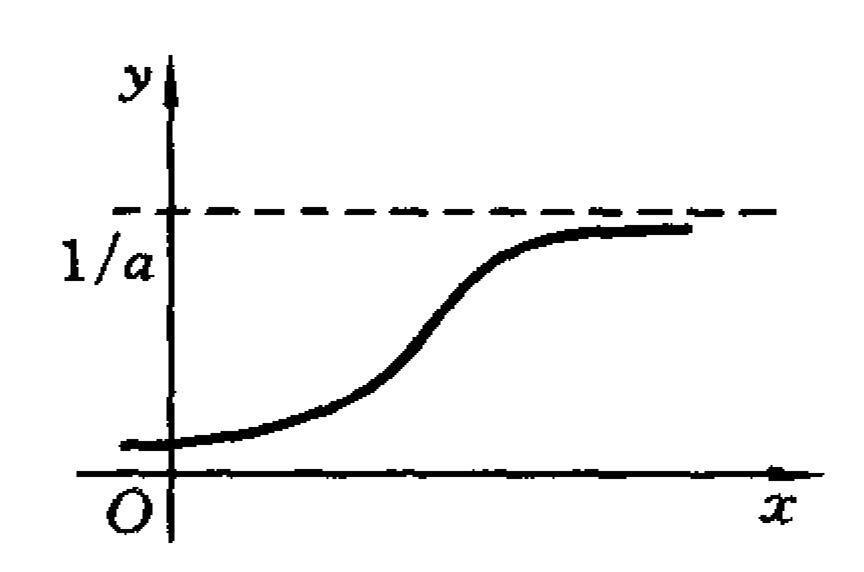


图 9.9

$$v=a+bu,$$

由数据计算 $\hat{a},\hat{b}$，即得

$$y=\frac{1}{\hat{a}+\hat{b}e^{-x}}.$$

其它如 $y=a+b\sin x$，只要令 $u=\sin x$ 即可化为直线方程.

三、多元线性回归简介

设随机变量 y 与 n 个普通变量 $x_1,x_2,\cdots,x_n$ 存在相关关系，并设

$$y=a+b_1x_1+b_2x_2+\cdots+b_mx_m+\varepsilon,\ \varepsilon\sim N(0,\sigma^2),$$

其中 $b_1,b_2,\cdots,b_m$ 为常数. 对变量 $x_1,x_2,\cdots,x_m$ 作 n 次观察，得观察值

$$(x_{i_1},x_{i_2},\cdots,x_{i_m};y_i),\ i=1,2,\cdots,n,$$

作为样本来估计参数 $a,b_1,b_2,\cdots,b_m$ 的估计值 $\hat{a},\hat{b}_1,\hat{b}_2,\cdots,\hat{b}_m$，求得 y 关于 $x_1,x_2,\cdots,x_n$ 的线性回归方程

$$\hat{y}=\hat{a}+\hat{b}_1x_1+\hat{b}_2x_2+\cdots+\hat{b}_mx_m.$$

1. 用极大似然法估计 $a,b_1,b_2,\cdots,b_m$

对样本的似然函数中的平方和

$$Q=\sum_{i=1}^{n}(y_i-\hat{y}_i)=\sum_{i=1}^{n}(y_i-a-b_1x_{i_1}-b_2x_{i_2}-\cdots-b_mx_{i_m})^2$$

解 Q 关于 a 和 $b_1,b_2,\cdots,b_m$ 的偏导数的方程组

$$\frac{\partial Q}{\partial a}=0,\quad \frac{\partial Q}{\partial b_1}=0,\quad \cdots,\quad \frac{\partial Q}{\partial b_m}=0,$$

解得$\hat{a},\hat{b}_1,\hat{b}_2,\cdots,\hat{b}_m$.

$$\hat{\boldsymbol{B}}=(\hat{a},\hat{b}_1\,\hat{b}_2\cdots\,\hat{b}_m)^{-1}=(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X})^{-1}X^{\mathrm{T}}Y,$$

其中
$$\boldsymbol{X}=\begin{pmatrix}1 & x_{11} & x_{12} & \cdots & x_{1m}\\ 1 & x_{21} & x_{22} & \cdots & x_{2m}\\ \vdots & \vdots & \vdots & & \vdots\\ 1 & x_{n1} & x_{n2} & \cdots & x_{nm}\end{pmatrix},\quad \boldsymbol{Y}=\begin{pmatrix}y_1\\ y_2\\ \vdots\\ y_n\end{pmatrix}.$$

若在显著性水平α下,拒绝

$$H_0:b_1=b_2=\cdots=b_m=0,$$

则认为回归的效果是显著的.

2. 未知参数σ^2的点估计

通常用$\hat{\sigma}=\frac{1}{n}Q(\hat{a},\hat{b}_1,\hat{b}_2,\cdots,\hat{b}_m)$,但$\hat{\sigma}^2$不是$\sigma^2$的无偏估计.

疑 难 解 析

1. 进行回归分析需满足哪些基本条件?

答 进行回归分析,对参数性检验对象要求必须满足以下三个基本条件:

(1) 正态性 被检验的对象或者因变量必须是服从正态分布$N(\mu,\sigma^2)$的随机变量.

(2) 方差齐性 被检验的各个总体的方差,应该是相等的.

(3) 独立性 对被检验的各对观察数据而言,从概率意义上应理解为是独立取得的.

处理实际问题时,往往不能事先预知这三条是否满足.诚如上节所提到的,正态性可由大数定律与中心极限定理来确定,方差齐性的检验用F检验来进行,而独立性一般凭实际经验判断.一般有一个近似结论就可以进行回归分析了.

2. 回归分析与相关分析有什么联系与区别?是否可以用同一种方法来研究?

答 相关关系是一种不确定性关系，也可以分为自变量与因变量加以考察. 因变量一般取可以测量的随机变量，而自变量在许多情况下不是随机的，是可控制的普通变量. 它表现为因变量的取值随自变量的变化而呈现一定的统计规律性，即当自变量取值确定时，因变量的取值也有确定的概率分布与之对应.

相关分析与回归分析的关系与区别表现在：相关分析一般是研究随机变量与随机变量之间的相关关系的，而回归分析研究随机变量与非随机变量之间的相关关系. 两者所使用的概念、理论和方法有所不同，得到的结果含义也不相同，但结果的形式却几乎完全一致. 因此，从应用与计算角度看，两者没有必要加以严格区别. 由于回归分析在数学处理上更为简便，因而不论自变量如何，都可当作非随机的普通变量看待，用回归分析方法研究变量间的相关关系.

3. 回归系数的最小二乘估计与极大似然估计有什么不同？它们的结果是否相同？

答 以一元正态线性回归为例. 其模型为

$$y_i=a+bx+\varepsilon_i,\quad i=1,2,\cdots,n,$$

其中

$$E(\varepsilon_i)=0.$$

最小二乘估计是指对 x,y 的 n 对试验值 $(x_i,y_i)(i=1,2,\cdots,n)$，作离差平方和

$$Q=\sum_{i=1}^{n}(y_i-a-bx_i)^2,$$

利用微积分中的极值方法，求出使 $Q(\hat{a},\hat{b})=\min Q(a,b)$ 的 $\hat{a}$ 和 $\hat{b}$，则 $\hat{y}=\hat{a}+\hat{b}x$ 为所求线性回归方程.

极大似然估计是由 $y_i=a+bx_i+\varepsilon_i$，$\varepsilon_i\sim N(0,\sigma^2)$，则 $y_i\sim N(a+bx_i,\sigma^2)$，得样本的极大似然函数

$$L=\left(\frac{1}{\sigma\sqrt{2\pi}}\right)^n\exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i-a-bx_i)^2\right].$$

要使 L 取最大值，则应使 $\sum_{i=1}^{n}(y_i-a-bx_i)^2$ 为最小. 用与最小二乘估计中同样的极值方法，求得使

$$Q=\sum_{i=1}^{n}(y_i-a-bx_i)^2$$

最小的 $\hat{a}$ 和 $\hat{b}$，得线性回归方程 $\hat{y}=\hat{a}+\hat{b}x$.

由于两种方法讨论的都是 $Q=\sum_{i=1}^{n}(y_i-a-bx_i)^2$，且由此求得 $\hat{a}$ 和 $\hat{b}$，故最小二乘估计与极大似然估计的结果是相同的.

4. 在线性回归确定后，影响预测精度的主要因素有哪些？

答 当线性回归方程 $\hat{y}=\hat{a}+\hat{b}x$ 已经确定，并经检验确认回归显著，则对给定的 x_0，y_0，置信度为 $1-\alpha$ 的预测区间为

$$\left(\hat{y}_0\mp t_{\alpha/2}(n-2)\hat{\sigma}\sqrt{1+1/n+(x_0-\bar{x})^2/l_{xx}}\right),$$

其中

$$l_{xx}=\sum_{i=1}^{n}(x_i-\bar{x})^2.$$

由此可知，影响预测精度的主要因素为：

(1) σ^2. 一般，σ^2 越小，精度越高.

(2) n. n 越大，精度越高，所以应尽量扩大样本容量.

(3) 自变量的取值 x_i. x_i 应尽量避免过于集中，但预测点 x_0 离 $\bar{x}$ 越近时精度越高.

5. 非线性回归的线性化过程是怎样进行的？关键是什么？

答 当具有相关关系的两个变量 y 和 x，不具有线性相关关系，而具有某种曲线相关关系时，可以通过适当的变量代换，将变量间的关系化为线性的形式，即通过必要的变量转换对它作线性化处理. 在两个变量回归的条件下，一般常用的是倒代换与对数代换，使曲线问题化为代换后的直线问题. 所处理的自变量与因变量间的关系，可以是双曲函数、幂函数、指数函数、负指数函数、对数函数等. 更复杂的曲线回归，将利用后面的多元线性回归方法来解决.

正确选择曲线类型，是正确地进行变量转换的前提，而正确的转换关系又是提高曲线回归精确度的根本. 由散点图的形状选择的线性化转换，往往不能一次就选准. 为准确起见，不妨同时作几种曲线加以比较.

方法、技巧与典型例题分析

一、一元线性回归问题

一元线性回归的解题过程是：(1) 建立观测结果 (x_1, y_1)，(x_2, y_2)，…，(x_n, y_n) 的散点图，确定 y 对 x 的相依关系的特点；(2) 估计回归系数 a，b 和方差 σ^2；(3) 检验回归方程的回归效果是否显著；(4) 利用回归方程进行预测与控制. 在计算数据过程中，要注意利用数据处理的一些技巧，以减小计算工作量.

例 1 某工业部门为了分析该部门的产量(单位：千件)x 与生产费用(单位：千元)y 之间的关系，随机抽取了 10 个企业作样本，得到数据如表 9.43 所示. 试建立 x 与 y 之间的回归方程式.

表 9.43

x	40	42	48	55	65	79	88	100	120	140
y	150	140	160	170	150	162	185	165	190	185

解 画散点图(略)，可以看出散点大致呈直线分布趋势. 设线性回归方程为 $y=a+bx$. 计算结果如表 9.44 所示.

计算数据如下：

$$l_{xx}=\sum_{i=1}^{n}(x_i-\overline{x})^2=\sum_{i=1}^{n}x_i^2-\frac{1}{n}\Big(\sum_{i=1}^{n}x_i\Big)^2$$

$$=70903-\frac{1}{10}\times 777^2=10530.1,$$

$$l_{yy}=\sum_{i=1}^{n}(y_i-\overline{y})^2=\sum_{i=1}^{n}y_i^2-\frac{1}{n}\Big(\sum_{i=1}^{n}y_i\Big)^2$$

$$=277119-\frac{1}{10}\times 1657^2=2554.1,$$

$$l_{xy}=\sum_{i=1}^{n}(x_i-\overline{x})(y_i-\overline{y})=\sum_{i=1}^{n}x_iy_i-\frac{1}{n}\left(\sum_{i=1}^{n}x_i\right)\left(\sum_{i=1}^{n}y_i\right)$$

$$=132938-\frac{1}{10}\times 777\times 1657=4189.1.$$

表 9.44

	x	y	x^2	xy	y^2
	40	150	1600	6000	22500
	42	140	1764	5880	19600
	48	160	2304	7680	25600
	55	170	3025	9350	28900
	65	150	4225	9750	22500
	79	162	6241	12798	26244
	88	185	7744	16280	34225
	100	165	10000	16500	27225
	120	190	14400	22800	36100
	140	185	19600	25900	34225
$\sum$	777	1657	70903	132938	277119

解得
$$\hat{b}=\frac{l_{xy}}{l_{xx}}=\frac{4189.1}{10530.1}=0.398.$$

$$\hat{a}=\frac{1}{n}\sum_{i=1}^{n}y_i-\hat{b}\overline{x}=165.7-0.398\times 77.7=134.78,$$

所以,线性回归方程为

$$\hat{y}=134.78+0.398x.$$

而σ^2的无偏估计为

$$\hat{\sigma}=\frac{1}{n-2}(l_{yy}-\hat{b}l_{xx})=\frac{1}{8}(2554.1-0.398\times 4189.1)$$
$$=110.85.$$

例 2 设关于某设备的使用年限 x 和所支出的维修费用(单

位：千元）y 如表 9.45 所示，求：

（1）y 关于 x 的回归方程，σ^2 的无偏估计；

（2）检验回归是否显著，并求 $x=7$ 时，维修费用 y 的 0.95 的预测区间.

表 9.45

x	2	3	4	5	6
y	2.2	3.8	5.5	6.5	7.0

解 （1）作散点图（略），数据呈直线分布趋势. 计算结果如表 9.46 所示.

表 9.46

	x	y	x^2	xy	y^2
	2	2.2	4	4.4	4.84
	3	3.8	9	11.4	14.44
	4	5.5	16	22.0	30.25
	5	6.5	25	32.5	42.25
	6	7.0	36	42.0	49.00
$\sum$	20	25	90	112.3	140.78

计算数据如下：

$$l_{xx}=\sum_{i=1}^{n}x_i^2-\frac{1}{n}\left(\sum_{i=1}^{n}x_i\right)^2=90-\frac{1}{5}\times 20^2=10,$$

$$l_{xy}=\sum_{i=1}^{n}x_iy_i-\frac{1}{n}\left(\sum_{i=1}^{n}x_i\right)\left(\sum_{i=1}^{n}y_i\right)=112.3-\frac{1}{5}\times 20\times 25=12.3,$$

$$l_{yy}=\sum_{i=1}^{n}y_i^2-\frac{1}{n}\left(\sum_{i=1}^{n}y_i\right)^2=140.78-\frac{1}{5}\times 25^2=15.78,$$

解得

$$\hat{b}=\frac{l_{xy}}{l_{xx}}=\frac{12.3}{10}=1.23,$$

$$\hat{a}=\frac{1}{n}\sum_{i=1}^{n}y_i-b\overline{x}=5-1.23\times 4=0.08.$$

所以，线性回归方程为

$$\hat{y}=0.08+1.23x.$$

σ^2 的无偏估计为

$$\hat{\sigma}^2=\frac{1}{n-2}(l_{yy}-\hat{b}l_{xy})=\frac{1}{3}(140.78-1.23\times 112.3)$$
$$=0.8837.$$

(2) 将 $x_0=7$ 代入回归方程得 $\hat{y}_0=8.69$. 因为 $n=5$，$t_{0.025}(3)=3.18$，所以 y_0 的置信度为 0.95 的置信区间为

$$\left(\hat{y}_0\pm t_{\alpha/2}(n-2)\hat{\sigma}\sqrt{1+1/n+(x_0-\bar{x})^2/l_{xx}}\right)$$
$$=(8.69\pm 3.18\times 0.94\times 1.45)=(4.3557,13.0243).$$

计算统计量 T，得

$$t=\frac{\hat{b}}{\hat{\sigma}}\sqrt{l_{xx}}=\frac{1.23}{\sqrt{0.8837}}\times\sqrt{10}=4.1376.$$

因为 $t_{0.025}(3)=3.1824<t=4.1376$，故知回归效果是显著的.

例 3 为研究某一化学反应过程中，温度(单位：C)x 对产品得率(质量分数，%)y 的影响，测得数据如表 9.47 所示. 求：

(1) 线性回归方程和 σ^2 的无偏估计；

(2) 检验回归效果是否显著和求 b 的置信区间($\alpha=0.05$).

表 9.47

x	100	110	120	130	140	150	160	170	180	190
y	45	51	54	61	66	70	74	78	85	89

解 (1) 画散点图(略)，知数据呈直线分布趋势.

计算数据如下：

$$\sum_{i=1}^{10}x_i=1450,\quad \sum_{i=1}^{10}y_i=673,\quad \sum_{i=1}^{10}x_i^2=218500,$$
$$\sum_{i=1}^{10}x_iy_i=101570,\quad \sum_{i=1}^{10}y_i^2=47225,\quad n=10.$$
$$l_{xx}=218500-\frac{1}{10}\times 1450^2=8250,$$

$$l_{xy}=101570-\frac{1}{10}\times1450\times673=3985,$$

$$l_{yy}=47225-\frac{1}{10}\times673^2=1932.1,$$

得
$$\hat{b}=l_{xy}/l_{xx}=0.4830,$$

$$\hat{a}=67.3-145\times0.4830=-2.7393.$$

所以,线性回归方程为

$$\hat{y}=-2.7393+0.4830x,$$

σ^2 的无偏估计为

$$\hat{\sigma}^2=\frac{1}{n-2}(l_{yy}-\hat{b}l_{xy})=0.9181.$$

(2) 因为 $t_{0.025}(8)=2.3060,\sigma=0.9581$,所以

$$|t|=\frac{|\hat{b}|}{\hat{\sigma}}\sqrt{l_{xx}}=\frac{10.48301}{\sqrt{0.9181}}\times\sqrt{8250}=45.7856.$$

由于 $|t|=45.7856>t_{0.025}(8)=2.3060$,故认为回归效果是显著的.

b 的置信度为 $1-\alpha$ 的置信区间为

$$(\hat{b}\mp t_{\alpha/2}(n-2)\cdot\hat{\sigma}/\sqrt{l_{xx}})$$
$$=(0.4830\mp2.3060\times0.9581/90.83)$$
$$=(0.4587,0.5073).$$

例 4 某地区第一年到第六年间每年的用电量(单位:10^8 kW·h)y 与年次 x 的统计数据如表9.48所示.求 y 对 x 的回归方程,并在 $\alpha=0.01$ 下作显著性检验.若该地区第七至第八年经济发展速度不变,试对第八年的用电量在 $\alpha=0.05$ 下进行预测.

表 9.48

x	1	2	3	4	5	6
y	10.4	11.4	13.1	14.2	14.8	15.7

解 作散点图(略),知数据分布呈直线趋势,所以,设

$$y=a+bx.$$

计算数据如下：

$$\sum_{i=1}^{6}x_i=21,\quad \sum_{i=1}^{6}y_i=79.6,$$

$$\bar{x}=3.5,\quad \bar{y}=13.27,$$

$$\sum_{i=1}^{6}x_i^2=91,\quad \sum_{i=1}^{6}y_i^2=1076.9,\quad \sum_{i=1}^{6}x_iy_i=297.5,$$

$$\frac{1}{6}\Big(\sum_{i=1}^{6}x_i\Big)^2=73.5,\quad \frac{1}{6}\Big(\sum_{i=1}^{6}y_i\Big)^2=1056.03,$$

$$\frac{1}{6}\Big(\sum_{i=1}^{6}x_i\Big)\Big(\sum_{i=1}^{6}y_i\Big)=278.6.$$

$$l_{xx}=17.5,\quad l_{xy}=18.9,\quad l_{yy}=20.87,$$

由公式得 $$\hat{b}=l_{xy}/l_{xx}=18.9/17.5=1.08,$$

$$\hat{a}=\frac{1}{6}\times79.6-1.08\times\frac{1}{6}\times21=9.49.$$

所以线性回归方程为

$$\hat{y}=9.49+1.08x,$$

σ^2 的无偏估计为

$$\hat{\sigma}^2=\frac{1}{4}(20.87-1.08\times18.9)=0.1145,\quad \hat{\sigma}=0.3384.$$

于是，由 $$|t|=\frac{\hat{b}}{\hat{\sigma}}\sqrt{l_{xx}}=13.351>t_{0.05}(4)=2.1318$$

知回归效果是显著的.

将 $x_0=8$ 代入线性回归方程，得第八年用电量 y_0 的置信度为 0.95 的置信区间为

$$(18.13\mp t_{0.025}(4)\times0.3384\sqrt{1+1/6+20.25/17.5})$$

$$=(16.70,19.56).$$

例5 表9.49 列出退火温度（单位：℃）x 对黄铜延展性 y 效应的试验数据，y 是以伸长率（%）计算的. 又设 x,y 均为正态变量，其方差与 x 无关，求 y 对于 x 的线性回归方程和 σ^2 的无偏估计.

表 9.49

x	300	400	500	600	700	800
y	40	50	55	60	67	70

解 画散点图(略),知数据呈直线分布趋势,所以,设

$$y=a+bx.$$

计算数据如下:

$$\sum_{i=1}^{6}x_i=3300,\quad \sum_{i=1}^{6}x_i^2=1990000,\quad \sum_{i=1}^{6}y_i=342,$$

$$\sum_{i=1}^{6}y_i^2=20114,\quad \sum_{i=1}^{6}x_iy_i=198400,$$

$$l_{xx}=17500,\quad l_{xy}=10300,\quad l_{yy}=620.$$

由公式得 $$\hat{b}=l_{xy}/l_{xx}=0.05886,$$

$$\hat{a}=\frac{1}{6}\times 342-\frac{1}{6}\times 550\times 0.05886=24.6287.$$

所以线性回归方程为

$$\hat{y}=24.6287+0.05886x,$$

σ^2 的无偏估计为

$$\hat{\sigma}^2=\frac{1}{4}(620-0.05886\times 10300)=3.4355.$$

例 6 设儿子身高 y 与父亲的身高 x 适合一元正态线性回归模型,观察了 10 对英国父子的身高(单位:in,1 in=25.4 mm),得数据如表 9.50 所示.

(1) 试建立 y 关于 x 的回归方程;

(2) 在 $\alpha=0.05$ 下对方程作显著性检验;

(3) 当 $x_0=69$ 时,求 y_0 的置信度为 0.95 的预测区间.

表 9.50

x	60	62	64	65	66	67	68	70	72	74
y	63.6	65.5	66	65.6	66.9	67.1	67.4	63.3	70.1	70

解 (1) 设回归方程为 $y=ax+b$,按所给数据计算,得

$$\sum_{i=1}^{10} x_i = 668, \quad \bar{x} = 66.8, \quad \sum_{i=1}^{10} x_i^2 = 44794,$$

$$\sum_{i=1}^{10} y_i = 665.1, \quad \bar{y} = 66.51, \quad \sum_{i=1}^{10} y_i^2 = 44283.93,$$

$$\sum_{i=1}^{10} x_i y_i = 44492.4,$$

$$l_{xx} = \sum_{i=1}^{10} x_i^2 - 10\bar{x}^2 = 171.6,$$

$$l_{yy} = \sum_{i=1}^{10} y_i^2 - 10\bar{y}^2 = 48.129,$$

$$l_{xy} = \sum_{i=1}^{10} x_i y_i - 10\bar{x}\bar{y} = 63.72,$$

所以 $\hat{b} = l_{xy}/l_{xx} = 0.3713, \quad \hat{a} = \bar{y} - \hat{b}\bar{x} = 41.7072.$

线性回归方程为

$$\hat{y} = 41.7072 + 0.3713x.$$

(2) 待检假设 $H_0: b = 0$. 因为

$$\hat{\sigma}^2 = \frac{1}{n-2}(l_{yy} - \hat{b}l_{xy}) = \frac{1}{8}(48.129 - 0.3713 \times 63.72)$$
$$= 24.4698/8 = 3.0587,$$

所以 $|t| = \frac{\hat{b}}{\hat{\sigma}}\sqrt{l_{xx}} = \frac{0.3713}{\sqrt{3.0587}} \times \sqrt{171.6} = 2.7811.$

又因为 $t_{0.025}(8) = 2.3060 < |t| = 2.7811$，所以拒绝 H_0，认为回归的效果是显著的.

也可用 F 统计量检验.

(3) y_0 的置信度为 $1-\alpha$ 的预测区间为

$$(\hat{y}_0 \mp t_{\alpha/2}(n-2)\sigma\sqrt{1 + 1/n + (\bar{x} - x_0)^2/l_{xx}}).$$

当 $x_0 = 69$ 时，$y_0 = 67.3269$，$t_{0.025}(8) = 2.3060$，故 y_0 的置信度为 0.95的预测区间为(63.0432,71.6106).

例7 如果两个变量 x, y 存在相关关系，其中 y 的值是难以测量的，而 x 的值是易于测得的，则可以根据 x 的测量值利用回归方

程 $\hat{y}=\hat{a}+\hat{b}x$ 去估计 y 的值，表9.51给出18个5～8岁儿童的重量(易测的，单位:kg)和体积(难测的，单位:$10^{-3}\,m^3$)．设 x,y 是正态变量，方差与 x 无关．求：

(1) y 关于 x 的线性回归方程 $\hat{y}=\hat{a}+\hat{b}x$；

(2) $x=14$ 时，y 的置信度为0.95的预测区间．

表 9.51

x	17.1	10.5	13.8	15.7	11.9	10.4	15.0	16.0	17.8
y	16.7	10.4	13.5	15.7	11.6	10.2	14.5	15.8	17.6
x	15.8	15.1	12.1	18.4	17.1	16.7	16.5	15.1	15.1
y	15.2	14.8	11.9	18.3	16.7	16.6	15.9	15.1	14.5

解 (1) 计算数据如下：

$$\sum_{i=1}^{18}x_i=270.1,\quad \bar{x}=15.006,\quad \sum_{i=1}^{18}x_i^2=4149.39,$$

$$\sum_{i=1}^{18}y_i=265,\quad \bar{y}=14.722,\quad \sum_{i=1}^{18}y_i^2=3996.14,$$

$$\sum_{i=1}^{18}x_iy_i=4071.71,$$

$$l_{xx}=\sum_{i=1}^{n}x_i^2-\frac{1}{n}\Big(\sum_{i=1}^{n}x_i\Big)^2=96.3894,$$

$$l_{yy}=\sum_{i=1}^{n}y_i^2-\frac{1}{n}\Big(\sum_{i=1}^{n}y_i\Big)^2=94.7511,$$

$$l_{xy}=\sum_{i=1}^{n}x_iy_i-\frac{1}{n}\Big(\sum_{i=1}^{n}x_i\Big)\Big(\sum_{i=1}^{n}y_i\Big)=95.2378,$$

依公式得 $\hat{b}=l_{xy}/l_{xx}=0.9881$，$\hat{a}=\bar{y}-\hat{b}\bar{x}=-0.1040$，所以，$y$ 关于 x 的线性回归方程为

$$\hat{y}=-0.1040+0.9881x.$$

(2) 因为 σ 的无偏估计为

$$\hat{\sigma}^2=\frac{1}{n-2}(l_{yy}-\hat{b}l_{xy})=\frac{1}{16}(94.7511-0.9881\times 95.2378)$$

$=0.0404.$

当 $x_0=14$ 时，$\hat{y}_0=-0.1040+0.9881\times14=13.7294$，所以，$x=14$ 时，y_0 的置信度为0.95 的预测区间为

$$(\hat{y}_0\mp t_{0.025}(16)\times\sqrt{0.0404}\times\sqrt{1+1/8+(14-15.005)^2/96.3894})$$

$$=(14.696,14.748).$$

例8 在服装标准的制定过程中，需调查获取一系列的数据．表9.52给出一组女青年的身高 x 与裤长 y 的数据(单位:cm)，求：

(1) 裤长 y 对身高 x 的回归方程；

(2) 在 $\alpha=0.01$ 下检验回归方程的显著性．

表9.52

x	168	162	160	160	156	157	159	168	159	162	158	156	165	158	166
y	107	103	103	102	100	100	101	107	100	102	100	99	105	101	105
x	162	150	152	156	159	156	164	168	165	162	158	157	172	147	155
y	105	97	98	101	103	99	107	108	106	103	101	101	110	95	99

解 (1) 计算数据，得

$$\sum_{i=1}^{30}x_i=4797,\quad \bar{x}=159.9,\quad \sum_{i=1}^{30}x_i^2=767949,$$

$$\sum_{i=1}^{30}y_i=3068,\quad \bar{y}=102.3,\quad \sum_{i=1}^{30}y_i^2=314112,$$

$$\sum_{i=1}^{30}x_iy_i=491124,$$

$$l_{xx}=\sum_{i=1}^{30}x_i^2-\frac{1}{30}\Big(\sum_{i=1}^{30}x_i\Big)^2=767949-\frac{1}{30}\times4797^2=908.7,$$

$$l_{xy}=\sum_{i=1}^{30}x_iy_i-\frac{1}{30}\Big(\sum_{i=1}^{30}x_i\Big)\Big(\sum_{i=1}^{30}y_i\Big)=550.8,$$

$$l_{yy}=\sum_{i=1}^{30}y_i^2-\frac{1}{30}\Big(\sum_{i=1}^{30}y_i\Big)^2=357.87.$$

依公式得 $\hat{b}=l_{xy}/l_{xx}=0.61$，$\hat{a}=\bar{y}-\hat{b}\bar{x}=4.76$，

所以，裤长 y 关于身高 x 的线性回归方程为

$$\hat{y}=4.76+0.61x.$$

(2) 用 F 检验法，$F=\frac{U}{Q}(n-2)$，待检假设 $H_0:b=0$.

$$U=\hat{b}l_{xy}=0.61\times550.8=335.99,$$

$$Q=l_{yy}-U=357.87-335.99=21.88,$$

$$F=(n-2)U/Q=28\times335.99/21.88=429.96.$$

而 $F_{0.99}(1,28)=7.64<F=429.96$，故拒绝 H_0，认为线性关系高度显著.

二、可化为线性回归的非线性回归问题

可化为线性回归的非线性回归问题，关键在于确定回归曲线的类型. 在一些不易确定的情况，可选择几种曲线加以比较，选择效果较好的一种.

例 9 某商品的需求量(单位:件) y 与价格(单位:元) x 的统计资料如表 9.53 所示，求需求函数的回归方程.

表 9.53

y	543	580	618	695	724	812	887	991	1186	1940
x	61	54	50	43	38	36	28	23	19	10

解 画散点图(略)，根据散点图选择 $y=ax^{-b}$ 来描绘需求量 y 与价格 x 的关系. 经变换，得

$$Z=\ln y=\ln a-b\ln x=\alpha+\beta t.$$

利用最小二乘法求得 α 和 β 的估计值

$$\hat{\alpha}=9.1206,\quad \hat{\beta}=-0.6902,$$

所以 $\hat{a}=e^{\hat{\alpha}}=9141.685,\quad \hat{b}=-\hat{\beta}=0.6902.$

故需求回归方程为

$$\hat{y}=9141.685x^{-0.6902}.$$

将 y 与 $\hat{y}$ 的值加以比较，结果如表 9.54 所示，可见 y 与 $\hat{y}$ 数据相近，效果较好.

表 9.54

y	543	580	618	695	724	812	887	991	1186	1940
$\hat{y}$	536	583	614	682	742	771	917	1050	1198	1886

例 10 为研究某企业的生产率(单位:件/周)x 与废品率(%)y 的关系,调查记录的数据如表 9.55 所示,试根据数据拟合出合适的曲线模型.

表 9.55

x	1000	2000	3000	3500	4000	4500	5000
y	5.2	6.5	6.8	8.1	10.2	10.3	13.0

解 画出散点图(略),观察数据显示趋势,既呈直线趋势,又呈曲线趋势(指数增长),故分别作直线与指数曲线拟合.

(1) 设 $y=ax+b$,计算数据如下:

$$\sum_{i=1}^{7}x_i=23000,\quad \sum_{i=1}^{7}x_i^2=87500000,\quad \bar{x}=3285.7143,$$

$$\sum_{i=1}^{7}y_i=60.1,\quad \sum_{i=1}^{7}y_i^2=560.27,\quad \bar{y}=8.5857,$$

$$\sum_{i=1}^{7}x_iy_i=219100,$$

$$l_{xx}=11928571.4286,\quad l_{yy}=516,\quad l_{xy}=21628.9.$$

依公式得 $\hat{b}=l_{xy}/l_{xx}=0.00181$,$\hat{a}=\bar{y}-\hat{b}\bar{x}=2.6386$,所以直线回归方程为

$$\hat{y}=2.6386+0.00181x.$$

(2) 设 $y=ae^{bx}$,取对数转化为 $\ln y=\ln a+bx$,得

$$\sum_{i=1}^{7}\ln y_i=14.75.$$

求得 $\ln a$ 和 b 的最小二乘估计为

$$\hat{b}=0.002,\quad \ln\hat{a}=1.3987,\quad \hat{a}=4.05,$$

所以指数曲线回归方程为

$$\hat{y}=4.05e^{0.002x}.$$

比较两个模型的残差平方和,以残差平方和小的为优.

直线回归模型的残差平方和$Q=5.3371$,指数曲线回归模型的残差平方和$Q=6.11$,故认为直线回归模型拟合程度更好.

例11 在彩色显影中,由经验知:形成染料光学密度y与析出银的光学密度x由公式

$$y=Ae^{b/x} \quad (b<0)$$

表示,测得试验数据如表9.56所示,求y关于x的回归方程.

表9.56

x	0.05	0.06	0.07	0.10	0.14	0.20	0.25	0.31	0.38	0.43	0.47
y	0.10	0.14	0.23	0.37	0.59	0.70	1.00	1.12	1.19	1.25	1.29

解 对公式$y=Ae^{b/x}$两边取对数,得

$$\ln y=\ln A+b/x,$$

化为

$$v=a+bu \quad (a=\ln A).$$

令$u_i=1/x,v_i=\ln y_i,i=1,2,\cdots,11$,得数据如表9.57所示.

表9.57

u	20.00	16.667	14.286	10.00	7.143	
v	−2.303	−1.966	−1.47	−0.994	−0.528	
u	5.00	4.00	3.226	2.632	2.326	2.128
v	−0.236	0	0.113	0.174	0.223	0.255

计算以下数据:

$$\bar{u}=7.946, \quad \bar{v}=-0.612,$$

$$l_{uu}=406.614, \quad l_{vv}=8.690, \quad l_{uv}=-59.343,$$

依公式得 $\hat{b}=l_{uv}/l_{uu}=-0.146, \quad \hat{a}=\bar{v}-\hat{b}\bar{u}=0.548.$

于是 $A=e^{0.548}=1.729,$

故回归方程为 $\hat{y}=1.729e^{-0.146/x}.$

例 12 我国在1981—1988年的八年间，全国居民人均年消费水平(单位:元)y和年份x的统计数据如表9.58所示. 以$t=1,2,\cdots,8$表示1981,1982,…,1988年度，试建立y对年度$t=(x-1980)$的经验回归方程.

表 9.58

x	1981	1982	1983	1984	1985	1986	1987	1988
y	249	267	289	329	406	451	513	643

解 由散点图(略)可以看出，数据分布呈指数曲线趋势. 试用函数

$$y=A+ae^{b(x-1980)}=240+ae^{bt},$$

其中$A=240$为任意选定的基点. 令$u=\ln(y-240)$，$\alpha=\ln a$，则得$u=\alpha+bt$.

计算以下数据：

$\bar{t}=4.625$，$\bar{u}=4.4932$，$\overline{t^2}=26.875$，$\overline{t\,u}=23.5575$，

依公式得$\hat{b}=0.5062$，$\hat{\alpha}=2.1520$，$\hat{a}=e^{\hat{\alpha}}=8.6020$，所以，经验回归方程为

$$\hat{u}=2.1520+0.5062t,$$

即
$$\hat{y}=240+8.6020e^{0.5062(x-1980)}.$$

例 13 表9.59所示的是美国旧轿车价格调查资料，其中x是轿车使用年数，y是相应的平均价格(单位:美元). 试建立y关于x的回归方程.

表 9.59

x	1	2	3	4	5	6	7	8	9	10
y	2651	1943	1494	1087	765	538	484	290	226	204

解 由散点图(略)看出，数据分布呈指数曲线趋势，故设

$$y=ce^{bx}.$$

两边取对数，得$\ln y=\ln c+bx$，化为

$$v=a+bx \quad (b>0,\ln c=a),$$

得数据如表 9.60 所示.

表 9.60

x	1	2	3	4	5	6	7	8	9	10
v	7.883	7.572	7.309	6.991	6.64	6.288	6.182	5.67	5.42	5.318

计算以下数据:

$$\bar{x}=5.5,\quad \bar{v}=6.5274,$$

$$l_{xx}=82.5,\quad l_{vv}=7.3663,\quad l_{xv}=-24.5586,$$

依公式得

$$\hat{b}=l_{xv}/l_{xx}=-0.2977,$$

$$\hat{a}=\bar{v}-\hat{b}\bar{x}=8.1646,\quad \hat{c}=e^{\hat{a}}=3514.28,$$

故 y 对 x 的回归方程为

$$\hat{y}=3514.28e^{-0.2977x}.$$

三、多元线性回归问题

多元线性回归问题比较复杂.计算过程由于涉及线性方程组的求解,可以用克莱姆法则或者矩阵形法.有时为了简便,还可以用编码形式表达不同的水平(见例 15).

例 14 某煤矿十年时间原煤生产的劳动生产率(单位:t/工)y、产量(单位:10^6 t)x_1 和掘进尺(单位:km)x_2 的统计数据如表 9.61所示.设 y 与 x_1,x_2 有相依关系 $y=a+b_1x_1+b_2x_2$,

(1) 建立 y 对 x_1,x_2 的经验回归方程;

(2) 求 σ^2 的无偏估计,并检验回归效果;

(3) 检验回归系数 b_1 与 b_2 的显著性.

表 9.61

年次	1	2	3	4	5	6	7	8	9	10
x_1	2.46	2.23	1.97	2.31	2.13	2.65	2.50	2.36	2.40	2.08
x_2	10.2	9.30	13.0	16.2	15.8	16.5	11.9	13.1	18.1	20.7
y	1.35	1.23	1.07	1.11	1.03	1.12	1.12	1.24	1.20	1.04

解 (1) 计算以下数据:

$$\sum_{i=1}^{10} x_{1i}=23.09,\quad \sum_{i=1}^{10} x_{1i}^2=53.707,\quad \sum_{i=1}^{10} x_{2i}=144.8,$$

$$\sum_{i=1}^{10} x_{2i}^2=2213.18,\quad \sum_{i=1}^{10} y_i=11.51,\quad \sum_{i=1}^{10} y_i^2=13.3413,$$

$$\sum_{i=1}^{10} x_{1i}x_{2i}=333.40,\quad \sum_{i=1}^{10} x_{2i}y_i=26.667,\quad \sum_{i=1}^{10} x_{2i}y_i=164.68,$$

$$l_{11}=5.3707-2.309^2=0.0392,$$

$$l_{22}=2.21318-14.48^2=11.6476,$$

$$l_{12}=33.34-2.309\times 14.48=-0.094,$$

$$l_{21}=l_{12}=-0.094,$$

$$l_{1y}=2.6667-2.309\times 1.151=0.0091,$$

$$l_{2y}=16.468-14.48\times 1.151=-0.199.$$

依公式得 $$\hat{a}=\bar{y}-\bar{x}\,\hat{b}_1-\bar{x}_2\hat{b}_2=0.9267,$$

$$\hat{b}_1=\frac{l_{1y}l_{22}-l_{2y}l_{12}}{l_{11}l_{22}-l_{12}^2}=0.1944,$$

$$\hat{b}_2=\frac{l_{2y}l_{11}-l_{1y}l_{21}}{l_{11}l_{22}-l_{12}^2}=-0.0155,$$

所以,经验回归方程为

$$\hat{y}=0.9267+0.1944x_1-0.0155x_2.$$

(2) 计算 y_i 的回归值,得数据如表 9.62 所示.

表 9.62

y_i	1.35	1.23	1.07	1.11	1.03	1.12	1.12	1.24	1.2	1.04
$\hat{y}_i$	1.259	1.227	1.118	1.136	1.107	1.199	1.241	1.194	1.125	1.021

$$Q_l=\sum_{i=1}^{10}(\hat{y}_i-y_i)^2=0.0463,\quad S_l^2=\frac{1}{7}\times 0.0463=0.0066,$$

$$Q_R=\sum_{i=1}^{10}(\hat{y}_i-\bar{y})^2=0.0507,\quad S_R^2=\frac{1}{2}\times 0.0507=0.0254,$$

故 σ^2 的无偏估计

$$S_l^2=\frac{1}{n-(k+1)}\sum_{i=1}^{10}(\hat{y}_i-y_i)^2=0.0066.$$

（3）因为统计量

$$F=\frac{Q_R/2}{Q_l/7}=\frac{0.0254}{0.0066}=3.85,$$

而$F_{0.10}(2,7)=3.26<3.85=F$，拒绝$H_0$，认为在水平$\alpha=0.1$下，回归效果是显著的.

（4）因为统计量

$$t_1=\hat{b}_1/(S_l\sqrt{c_{11}})=\frac{0.1944}{0.0812\sqrt{2.6}}=1.485,$$

$$t_2=\hat{b}_2/(S_l\sqrt{c_{22}})=\frac{0.0155}{0.0812\sqrt{0.0088}}=2.035,$$

（c_{11},c_{22}是矩阵$(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X})^{-1}$的元素），查表知

$$t_{0.1}(7)=1.89<t_2=2.035,\quad t_{0.2}(7)=1.41<t_1=1.485,$$

故认为x_1在$\alpha=0.2$下显著，x_2在$\alpha=0.1$下显著.

若用矩阵运算求回归方程，则有

$$\boldsymbol{Y}=\begin{pmatrix}1.35\\1.23\\1.04\end{pmatrix},\quad \boldsymbol{X}=\begin{pmatrix}1&2.46&10.2\\1&2.23&9.30\\1&2.08&20.7\end{pmatrix},\quad \hat{\boldsymbol{b}}=\begin{pmatrix}\hat{a}\\\hat{b}_1\\\hat{b}_2\end{pmatrix},$$

$$\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}=\begin{pmatrix}10&23.09&144.8\\23.09&53.761&333.404\\144.8&333.404&2213.18\end{pmatrix},\quad \boldsymbol{X}^{\mathrm{T}}\boldsymbol{Y}=\begin{pmatrix}11.51\\26.667\\164.675\end{pmatrix},$$

$$|\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}|=447.8698,$$

$$(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X})^{-1}=\begin{pmatrix}17.2032&-6.3086&-1.752\\-6.3086&2.6007&-0.021\\-1.752&-0.021&0.0088\end{pmatrix},$$

得

$$\hat{\boldsymbol{b}}=\begin{pmatrix}\hat{a}\\\hat{b}_1\\\hat{b}_2\end{pmatrix}=(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X})^{-1}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{Y}=\begin{pmatrix}0.9267\\0.1944\\-0.0155\end{pmatrix}.$$

与(1)的结果相同.

例 15 某种化工产品的得率 y 与反应温度 x_1、反应时间 x_2 及反应物浓度 x_3 有关. 设对于给定的 x_1,x_2,x_3，得率 y(%)服从正态分布，且方差与 x_1,x_2,x_3 无关. 今得试验结果如表 9.63 所示，其中 x_1,x_2,x_3 均为二水平且均以编码形式表达.

(1) 设 $\mu(x_1,x_2,x_3)=a+b_1x_1+b_2x_2+b_3x_3$，求 y 的多元线性回归方程.

(2) 若认为反应时间不影响得率，即 $\mu(x_1,x_2,x_3)=\beta_0+\beta_1x_1+\beta_2x_2$，求 y 的多元线性回归方程.

表 9.63

x_1	−1	−1	−1	−1	1	1	1	1
x_2	−1	−1	1	1	−1	−1	1	1
x_3	−1	1	−1	1	−1	1	−1	1
得率 y	7.6	10.3	9.2	10.2	8.4	11.1	9.8	12.6

解 (1) 由所给数据写出矩阵

$$\boldsymbol{X}=\begin{pmatrix}1 & -1 & -1 & -1\\1 & -1 & -1 & 1\\1 & -1 & 1 & -1\\1 & -1 & 1 & 1\\1 & 1 & -1 & -1\\1 & 1 & -1 & 1\\1 & 1 & 1 & -1\\1 & 1 & 1 & 1\end{pmatrix},\quad \boldsymbol{Y}=\begin{pmatrix}7.6\\10.3\\9.2\\10.2\\8.4\\11.1\\9.8\\12.6\end{pmatrix},\quad \boldsymbol{B}=\begin{pmatrix}a\\b_1\\b_2\\b_3\end{pmatrix},$$

故 $$\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}=\begin{pmatrix}8 & 0 & 0 & 0\\0 & 8 & 0 & 0\\0 & 0 & 8 & 0\\0 & 0 & 0 & 8\end{pmatrix},\quad (\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X})^{-1}=\frac{1}{8}\begin{pmatrix}1 & 0 & 0 & 0\\0 & 1 & 0 & 0\\0 & 0 & 1 & 0\\0 & 0 & 0 & 1\end{pmatrix},$$

$$X^TY=\begin{pmatrix}79.2\\4.6\\4.4\\9.2\end{pmatrix},\quad B=(X^TX)^{-1}X^TY=\begin{pmatrix}9.9\\0.575\\0.55\\1.15\end{pmatrix}.$$

所以，y 的三元线性回归方程为

$$\hat{y}=9.9+0.575x_1+0.55x_2+1.15x_3.$$

(2) 此时，只需略去 x_2 项，即得

$$\hat{y}=9.9+0.575x_1+1.15x_3.$$

例16 某公司在15个地区的某种商品的销量(单位：罗，1罗=144件)y、各地区人口数(单位：千人)x_1 和平均每户总收入(单位：元)x_2 如表9.64所示. 设 y 对 x_1,x_2 有线性相依关系，试建立 y 对 x_1,x_2 的经验回归方程，并求 σ^2 的无偏估计.

表 9.64

x_1	274	180	375	205	86	265	98	330
x_2	2450	3254	3802	2838	2347	3782	3008	2450
y	162	120	223	131	67	169	81	192
x_1	195	53	430	372	236	157	370	
x_2	2137	2560	4020	4427	2660	2088	2605	
y	116	55	252	232	144	103	212	

解 计算以下数据：

$$\sum_{i=1}^{15}x_{1i}=3626,\quad \sum_{i=1}^{15}x_{2i}=44428,\quad \sum_{i=1}^{15}x_{1i}x_{2i}=11419181,$$

$$\sum_{i=1}^{15}x_{1i}^2=1067614,\quad \sum_{i=1}^{15}x_{2i}^2=139063428,\quad \sum_{i=1}^{15}x_{1i}y_i=647107,$$

$$\sum_{i=1}^{15}y_i=2259,\quad \sum_{i=1}^{15}y_j^2=394107,\quad \sum_{i=1}^{15}x_{2i}y_i=7096619.$$

$$l_{11}=12739.25,\quad l_{22}=498241.5,\quad l_{12}=45296.86,$$

$$l_{21}=45296.86,\quad l_{1y}=6735.43,\quad l_{2y}=27050.83.$$

依公式得

$$\hat{b}_1=0.496,\quad \hat{b}_2=0.0092,\quad \hat{a}=3.453,$$

所以，y 对 x_1,x_2 的经验回归方程为

$$\hat{y}=3.453+0.496x_1+0.0092x_2.$$

y_i 的回归值如表 9.65 所示，于是算得

$$Q_l=\sum_{i=1}^{15}(y_i-\hat{y}_i)^2=56.884,\quad S_l^2=4.74,$$

所以，σ^2 的无偏估计值为 $S_l^2=4.74$.

表 9.65

y_i	162	120	223	131	67	169	81	192
$\hat{y}_i$	161.90	122.67	224.43	131.24	67.70	169.69	79.73	189.67
y_i	116	55	252	232	144	103	212	
$\hat{y}_i$	119.83	53.29	253.72	228.69	144.98	100.53	210.938	

例 17 一种合金在某种添加剂的不同浓度 x 之下，各做三次试验，得强度 y(单位：MPa)数据如表 9.66 所示. 以模型 $y=a+b_1x+b_2x^2+\varepsilon$, $\varepsilon\sim N(0,\sigma^2)$ 拟合数据，其中 a,b_1,b_2,σ^2 均与 x 无关，求回归方程 $\hat{y}=\hat{a}+\hat{b}_1x+\hat{b}_2x^2$.

表 9.66

x	10.0	15.0	20.0	25.0	30.0
y	25.2	29.8	31.2	31.7	29.4
	27.3	31.1	32.6	30.1	30.8
	28.7	27.8	29.7	32.3	32.8

解 令 $t_1=x,t_2=x^2$，模型化为

$$y=a+b_1t_1+b_2t_2.$$

$$X=\begin{pmatrix}1 & 10 & 100\\1 & 15 & 225\\1 & 20 & 400\\1 & 25 & 625\\1 & 30 & 900\end{pmatrix},\quad Y=\begin{pmatrix}27.067\\29.567\\31.167\\31.167\\31\end{pmatrix},\quad B=\begin{pmatrix}a\\b_1\\b_2\end{pmatrix},$$

$$X^{T}X=\begin{pmatrix}15 & 300 & 6750\\300 & 6750 & 165000\\6750 & 165000 & 4263750\end{pmatrix},\quad X^{T}Y=\begin{pmatrix}450.9\\9155\\207990\end{pmatrix},$$

$$(X^{T}X^{-1})=\frac{1}{\Delta}=\begin{pmatrix}1555312500 & -165375000 & 3937500\\-165375000 & 18393750 & -450000\\3937500 & -450000 & 112500\end{pmatrix},$$

其中
$$\Delta=295312000.$$

故
$$B=\begin{pmatrix}a\\b_1\\b_2\end{pmatrix}=(X^{T}X)^{-1}X^{T}Y=\begin{pmatrix}19.0336\\1.0086\\-0.0204\end{pmatrix}.$$

所以，y 的经验回归方程为

$$\hat{y}=19.0336+1.0086x-0.0204x^2.$$

例 18 养猪场为了估算猪的毛重（单位：kg）y 与其身长（单位：cm）x_1、肚围（单位：cm）x_2 的关系，测量了 14 头猪，所得数据如表 9.67 所示. 经验表明，y 与 x_1，x_2 存在线性相依关系，试求经验回归方程.

表 9.67

x_1	41	45	51	52	59	62	69	72	78	80	90	92	98	103
x_2	49	58	62	71	62	74	71	74	79	84	85	94	91	95
y	28	39	41	44	43	50	51	57	63	66	70	76	80	84

解 设线性回归方程为 $y=a+b_1x_1+b_2x_2$. 计算数据得

$$\bar{x}_1=\frac{1}{14}\sum_{i=1}^{14}x_{1i}=70.86,\quad \bar{x}_2=\frac{1}{14}\sum_{i=1}^{14}x_{2i}=74.93,$$

$$\bar{y}=\frac{1}{14}\sum_{i=1}^{14}y_i=56.57,$$

$$l_{11}=\sum_{i=1}^{14}x_{1i}^2-\frac{1}{14}\left(\sum_{i=1}^{14}x_{1i}\right)^2=5248.05,$$

$$l_{22}=\sum_{i=1}^{14}x_{2i}^2-\frac{1}{14}\left(\sum_{i=1}^{14}x_{2i}\right)^2=2549.93,$$

$$l_{12}=l_{21}=\sum_{i=1}^{14}x_{1i}x_{2i}-\frac{1}{14}\left(\sum_{i=1}^{14}x_{1i}\right)\left(\sum_{i=1}^{14}x_{2i}\right)=3495.44,$$

$$l_{10}=\sum_{i=1}^{14}x_{1i}y_i-\frac{1}{14}\left(\sum_{i=1}^{14}x_{1i}\right)\left(\sum_{i=1}^{14}y_i\right)=4400.30,$$

$$l_{20}=\sum_{i=1}^{14}x_{2i}y_i-\frac{1}{14}\left(\sum_{i=1}^{14}x_{2i}\right)\left(\sum_{i=1}^{14}y_i\right)=3036.72,$$

依公式得

$$\hat{b}_1=(l_{10}l_{22}-l_{12}l_{20})/(l_{11}l_{22}-l_{12}^2)=0.52,$$

$$\hat{b}_2=(l_{20}l_{11}-l_{21}l_{10})/(l_{11}l_{22}-l_{12}^2)=0.48,$$

$$\hat{a}=\bar{y}-\hat{b}_1\bar{x}_1-\hat{b}_2\bar{x}_2=-16.24,$$

所以，y 对 x_1，x_2 的线性回归方程为

$$\hat{y}=-16.24+0.52x_1+0.48x_2.$$

[General Information]
书名=概率论与数理统计 内容、方法与技巧
作者=孙清华，孙吴主编
页数=493
SS号=11726098
出版日期=2006年5月